BAYERISCHE JULIUS-MAXIMILIANS-UNIVERSITÄT WÜRZBURG

FAKULTÄT FÜR BIOLOGIE

LEHRSTUHL FÜR MIKROBIOLOGIE

The immune transcriptome and proteome of the ant *Camponotus floridanus* and vertical transmission of its bacterial endosymbiont *Blochmannia floridanus*

DISSERTATION

zur Erlangung des naturwissenschaftlichen Doktorgrades

der bayerischen Julius-Maximilians-Universität Würzburg

vorgelegt von

Maria Kupper

aus Lutherstadt Eisleben

Würzburg, September 2016

The immune transcriptome and proteome of the ant *Camponotus floridanus* and vertical transmission of its bacterial endosymbiont *Blochmannia floridanus*

vorgelegt von

Maria Kupper

aus Lutherstadt Eisleben

Würzburg, September 2016

Eingereicht am: ....

Mitglieder der Promotionskommission:

Vorsitzender: Prof. Dr. ....

Erstgutachter: Prof. Dr. Roy Gross

Zweitgutachter: Prof. Dr. Heike Feldhaar

Tag des Promotionskolloquiums: ....

Doktorurkunde ausgehändigt am: ....

Eidesstattliche Erklärungen

gemäß § 4 Abs. 3 Satz 3, 5 und 8 der Promotionsordnung für die Fakultät für Biologie der Julius-Maximilians-Universität Würzburg

Hiermit erkläre ich, dass die vorliegende Arbeit „The immune transcriptome and proteome of the ant *Camponotus floridanus* and vertical transmission of its bacterial endosymbiont *Blochmannia floridanus*" von mir selbständig und nur unter Zuhilfenahme der angegebenen Quellen und Hilfsmittel angefertigt wurde.

Diese Dissertation hat weder in gleicher noch in ähnlicher Form in einem anderen Prüfungsverfahren vorgelegen. Ich habe früher, außer den mit dem Zulassungsgesuch urkundlich vorgelegten Graden, keine weiteren akademischen Grade erworben oder zu erwerben versucht.

Würzburg,

_____

Maria Kupper

Acknowledgments

"But Dreams Are Meaningful When You Work Toward Them In The Real World."

(Major Motoko Kusanagi, GITS:SAC)

Table of Contents

Summary

The evolutionary success of insects is believed to be at least partially facilitated by symbioses between insects and prokaryotes. Bacterial endosymbionts confer various fitness advantages to their hosts, for example by providing nutrients lacking from the insects' diet thereby enabling the inhabitation of new ecological niches. The Florida carpenter ant *Camponotus floridanus* harbours endosymbiotic bacteria of the genus *Blochmannia*. These primary endosymbionts mainly reside in the cytoplasm of bacteriocytes, specialised cells interspersed into the midgut tissue, but they were also found in oocytes which allows their vertical transmission. The social lifestyle of *C. floridanus* may facilitate the rapid spread of infections amongst genetically closely related animals living in huge colonies. Therefore, the ants require an immune system to efficiently combat infections while maintaining a "chronic" infection with their endosymbionts.

In order to investigate the immune repertoire of the ants, the Illumina sequencing method was used. The previously published genome sequence of *C. floridanus* was functionally re-annotated and 0.53% of *C. floridanus* proteins were assigned to the gene ontology (GO) term subcategory "immune system process". Based on homology analyses, genes encoding 510 proteins with possible immune function were identified. These genes are involved in microbial recognition and immune signalling pathways but also in cellular defence mechanisms, such as phagocytosis and melanisation. The components of the major signalling pathways appear to be highly conserved and the analysis revealed an overall broad immune repertoire of the ants though the number of identified genes encoding pattern recognition receptors (PRRs) and antimicrobial peptides (AMPs) is comparatively low. Besides three genes coding for homologs of thioester-containing proteins (TEPs), which have been shown to act as opsonins promoting phagocytosis in other insects, six genes encoding the AMPs defesin-1 and defensin-2, hymenoptaecin, two tachystatin-like peptides and one crustin-like peptide are present in the ant genome. Although the low number of known AMPs in comparison to 13 AMPs in the honey bee *Apis mellifera* and 46 AMPs in the wasp *Nasonia vitripennis* may indicate a less potent immune system, measures summarised as external or social immunity may enhance the immune repertoire of *C. floridanus*, as it was discussed for other social insects. Also, the hymenoptaecin multipeptide precursor protein may be processed to yield seven possibly bioactive peptides. In this work, two hymenoptaecin derived peptides were heterologously expressed and purified. The preliminary antimicrobial activity assays indicate varying bacteriostatic effects of different hymenoptaecin derived peptides against *Escherichia coli* D31 and *Staphylococcus aureus* which suggests a functional amplification of the immune response further increasing the antimicrobial potency of the ants.

Furthermore, 257 genes were differentially expressed upon immune challenge of *C. floridanus* and most of the immune genes showing differential expression are involved in recognition of microbes or encode immune effectors rather than signalling components. Additionally, genes coding for proteins involved in storage and metabolism were downregulated upon immune challenge suggesting a trade-

off between two energy-intensive processes in order to enhance effectiveness of the immune response. The analysis of gene expression via qRT-PCR was used for validation of the transcriptome data and revealed stage-specific immune gene regulation. Though the same tendencies of regulation were observed in larvae and adults, expression of several immune-related genes was generally more strongly induced in larvae. Immune gene expression levels depending on the developmental stage of *C. floridanus* are in agreement with observations in other insects and might suggest that animals from different stages revert to individual combinations of external and internal immunity upon infection.

The haemolymph proteome of immune-challenged ants further established the immune-relevance of several proteins involved in classical immune signalling pathways, e.g. PRRs, extracellularly active proteases of the Toll signalling pathway and effector molecules such as AMPs, lysozymes and TEPs. Additionally, non-canonical proteins with putative immune function were enriched in immune-challenged haemolymph, e.g. Vitellogenins, NPC2-like proteins and Hemocytin. As known from previous studies, septic wounding also leads to the upregulation of genes involved in stress responses. In the haemolymph, proteins implicated in protein stabilisation and in the protection against oxidative stress and insecticides were enriched upon immune challenge. In order to identify additional putative immune effectors, haemolymph peptide samples from immune-challenged larvae and adults were analysed. The analysis in this work focussed on the identification of putative peptides produced via the secretory pathway as previously described for neuropeptides of *C. floridanus*. 567 regulated peptides derived from 39 proteins were identified in the larval haemolymph, whereas 342 regulated peptides derived from 13 proteins were found in the adult haemolymph. Most of the peptides are derived from hymenoptaecin or from putative uncharacterised proteins. One haemolymph peptide of immune-challenged larvae comprises the complete amino acid sequence of a predicted peptide derived from a Vitellogenin. Though the identified peptide lacks similarities to any known immune-related peptide, it is a suitable candidate for further functional analysis.

To establish a stable infection with the endosymbionts, the bacteria have to be transmitted to the next generation of the ants. The vertical transmission of *B. floridanus* is guaranteed by bacterial infestation of oocytes. This work presents the first comprehensive and detailed description of the localisation of the bacterial endosymbionts in *C. floridanus* ovaries during oogenesis. Whereas the most apical part of the germarium, which contains the germ-line stem cells, is not infected by the bacteria, small somatic cells in the outer layers of each ovariole were found to be infected in the lower germarium. Only with the beginning of cystocyte differentiation, endosymbionts are exclusively transported from follicle cells into the growing oocytes, while nurse cells were never infected with *B. floridanus*. This infestation of the oocytes by bacteria very likely involves exocytosis-endocytosis processes between follicle cells and the oocytes. A previous study suggested a down-modulation of the immune response in the midgut tissue which may promote endosymbiont tolerance. Therefore, the expression of several potentially relevant immune genes was analysed in the ovarial tissue by qRT-PCR. The relatively low

expression of genes involved in Toll and IMD signalling, and the high expression of genes encoding negative immune regulators, such as *PGRP-LB*, *PGRP-SC2,* and *tollip*, strongly suggest that a down-modulation of the immune response may also facilitate endosymbiont tolerance in the ovaries and thereby contribute to their vertical transmission.

Overall, the present thesis improves the knowledge about the immune repertoire of *C. floridanus* and provides new candidates for further functional analyses. Moreover, the involvement of the host immune system in maintaining a "chronic" infection with symbiotic bacteria was confirmed and extended to the ovaries.

Zusammenfassung

Der evolutionäre Erfolg von Insekten wird zumindest teilweise Symbiosen zwischen Insekten und Prokaryonten zugeschrieben. Dabei übertragen bakterielle Symbionten verschiedenste Fitnessvorteile an ihre Wirte. Beispielsweise ermöglicht die Bereitstellung von Nährstoffen, welche in der Nahrung des Insekts fehlen, die Erschließung neuer ökologischer Nischen. Die Florida Rossameise *Camponotus floridanus* trägt endosymbiontische Bakterien der Gattung *Blochmannia*. Diese primären Endosymbionten kommen hauptsächlich im Zytoplasma von spezialisierten Zellen des Mitteldarms, den sogenannten Bakteriozyten, vor. Blochmannien wurden aber auch in Oozyten und Eiern gefunden, was ihre vertikale Übertragung an Individuen der nächsten Generation ermöglicht. Als soziale Insekten leben *C. floridanus* in großen Kolonien von nah verwandten Individuen. Ihre Lebensweise begünstigt möglicherweise die schnelle Ausbreitung von Infektionen, weshalb erwartet werden müsste, dass die Ameisen ein effizientes Immunsystem besitzen. Gleichzeitig muss jedoch die „chronische" Infektion mit den bakteriellen Symbionten aufrechterhalten werden.

In der vorliegenden Arbeit wurde das Immunrepertoire der Ameisen mittels Illumina Sequenzierung charakterisiert. Zunächst wurde das vor kurzem publizierte Genom von *C. floridanus* funktionell re-annotiert. Dabei wurden 0.53% der annotierten Proteine der GO-Unterkategorie "Prozesse des Immunsystems" zugeordnet. Basierend auf Homologieanalysen wurden Gene identifiziert, die für 510 Immunproteine kodieren. Die Genprodukte spielen eine Rolle bei der Erkennung von Mikroben und in den Signalwegen des Immunsystems, sind jedoch auch an Prozessen der zellulären Immunantwort, wie beispielsweise Phagozytose und Melanisierung, beteiligt. Dabei sind Komponenten der Hauptsignalwege hoch konserviert. Obwohl die Anzahl der identifizierten Proteine, die Fremdorganismen erkennen (PRRs), und die Anzahl an antimikrobiellen Peptiden (AMPs) vergleichsweise gering ist, verfügt *C. floridanus* insgesamt über ein umfangreiches Immunrepertoire. Neben drei Genen, die für Thioester-enthaltende Proteine (TEPs) kodieren und wie in anderen Insekten möglicherweise eine Rolle als Opsonine bei der Phagozytose spielen, wurden sechs AMP-Gene identifiziert. Diese kodieren für Defesin-1 und Defensin-2, Hymenoptaecin, zwei Tachystatin-ähnliche und ein Crustin-ähnliches Peptid. Die geringe Anzahl an bekannten AMPs im Vergleich zur Honigbiene *Apis mellifera* (13 AMPs) und Wespe *Nasonia vitripennis* (46 AMPs) könnte ein möglicherweise geringeres Potential des Immunsystems anzeigen. Allerdings könnten zusätzliche Maßnahmen, die unter dem Begriff „Soziale Immunität" zusammengefasst werden, das Immunrepertoire von *C. floridanus* ergänzen, wie es schon für andere Insekten diskutiert wurde. Zudem könnten durch proteolytische Prozessierung des Hymenoptaecin Multipeptid Präkursormoleküls sieben mögliche antimikrobielle Peptide freigesetzt werden. Für die vorliegende Arbeit wurden zwei verschiedene dieser Hymenoptaecin Peptide heterolog exprimiert und aufgereinigt. Die vorläufige funktionelle Charakterisierung der Peptide zeigt, dass diese Peptide möglicherweise bakteriostatische Wirkung mit einem unterschiedlichen Wirkspektrum gegen

Zusammenfassung

Escherichia coli D31 und *Staphylococcus aureus* entfalten. Dies erlaubt die Annahme, dass die Expression des Hymenoptaecins zu einer funktionellen Amplifikation der Immunantwort führt und das Immunrepertoire der Ameisen erweitert.

Nach Injektion von bakteriellem Material in die Ameisen wurde die Expression von 257 Genen reguliert. Viele dieser Gene kodieren für Proteine zur Erkennung von Pathogenen oder kodieren für Effektoren des Immunsystems. Komponenten der Signalwege zeigten dagegen kaum Veränderungen in ihrer Expression auf. Außerdem zeigten Gene, die für Speicherproteine oder Proteine des Metabolismus kodieren, generell eine geringere Expression nach Stimulierung des Immunsystems auf. Dies lässt einen Ausgleich zwischen zwei energieintensiven Prozessen vermuten, um eine effektive Immunantwort zu ermöglichen. Darüber hinaus zeigt die Validierung der Expressionsdaten mittels qRT-PCR eine Abhängigkeit der Expression mehrerer Gene vom Entwicklungsstadium der Ameisen auf. Generell wurden die gleichen Tendenzen in der Regulation der Expression dieser Gene nach Immunstimulierung beobachtet. Allerdings wurde die Expression mehrerer immunrelevanter Gene in Larven weit stärker induziert als in Adulten. Wie es auch schon für andere Insekten gezeigt wurde, scheinen *C. floridanus* Larven und Arbeiterinnen auf individuelle Kombinationen externer und interner Immunfaktoren zurückzugreifen.

Die vorher beschriebenen Transkriptomdaten wurden durch die Charakterisierung des Hämolymph-Proteoms von *C. floridanus* nach Immunstimulation ergänzt, wodurch die Immunrelevanz vieler Faktoren auch auf Proteinebene bestätigt werden konnte. Beispielsweise wurden zahlreiche PRRs und extrazellulär aktive Proteasen des Toll-Signalwegs, aber auch Immuneffektoren wie AMPs, Lysozyme und TEPs in der Hämolymphe identifiziert. Zusätzlich führte die Immunstimulation in Larven und Adulten zur Anreicherung nicht-kanonischer Proteine mit möglicher Immunfunktion, beispielsweise Vitellogenine, NPC2-ähnliche Proteine und Hemocytin. Aus einer früheren Arbeit ist bekannt, dass septische Verwundungen zusätzlich die transkriptionelle Aktivierung von Genen der Stressantwort hervorrufen können. So wurden auch in der Hämolymphe Proteine entdeckt, die eine Rolle bei der Stabilisierung von Proteinen, und dem Schutz gegen oxidativen Stress und Insektizide spielen. Zur Identifizierung weiterer möglicher Peptideffektoren wurden Hämolymphpeptid-Proben von immunstimulierten Larven und Adulten analysiert. Der Fokus der Analyse lag dabei auf der Identifizierung von Peptiden, die auf dem sekretorischen Weg gebildet werden, wie es zuvor für Neuropeptide von *C. floridanus* beschrieben worden war. 567 differentiell regulierte Peptide, die von 39 Proteinen abstammen, wurden in Larvenhämolymphe identifiziert, wohingegen in der Hämolymphe von Adulttieren 342 derartige Peptide, die 13 Proteinen zugeordnet werden können, gefunden wurden. Die meisten dieser Peptide können Hymenoptaecin oder bisher noch nicht charakterisierte Proteinen zugeordnet werden. Jedoch wurde ein Peptid in larvaler Hämolymphe identifiziert, dessen Aminosäuresequenz vollständig mit der Sequenz eines vorhergesagten, von Vitellogenin stammenden Peptids übereinstimmt. Weil dieses Peptid keine Ähnlichkeiten zu anderen

bereits charakterisierten antimikrobiellen Peptiden aufweist, stellt es einen geeigneten Kandidaten für weitere funktionelle Analysen dar.

Die bakterielle Infektion von Oozyten ermöglicht die transovariale Übertragung von *B. floridanus* und ermöglicht damit die Etablierung einer stabilen Infektion in der nächsten Wirtsgeneration. Die vorliegende Arbeit beinhaltet die erste umfassende und detaillierte Beschreibung der Lokalisation bakterieller Endosymbionten in Ovarien von *C. floridanus*. Im apikalen Germarium, in welchem sich die Keimbahn-Stammzellen befinden, liegt noch keine bakterielle Infektion des Gewebes vor. In späteren Segmenten des Germariums jedoch können Blochmannien das erste Mal in kleinen somatischen Zellen der äußeren Schicht jeder Ovariole detektiert werden. Mit beginnender Zystozytendifferenzierung werden die Endosymbionten von Follikelzellen ausschließlich in die heranwachsenden Oozyten transportiert, wobei sehr wahrscheinlich Exozytose-Endozytose-Prozesse involviert sind. Nährzellen zeigen zu keinem Zeitpunkt während der Oogenese eine bakterielle Infektion auf. Da in einer früheren Studie vorgeschlagen wurde, dass eine signifikant reduzierte Anregung der Immunantwort im Mitteldarmgewebe zur Toleranz der Endosymbionten beitragen könnte, wurde auch die Expression ausgewählter Immungene in den Ovarien durch qRT-PCR untersucht. Die relativ geringe Expression von Genen des Toll- und des IMD-Signalwegs und die zusätzlich vergleichsweise starke Genexpression negativer Regulatoren des Immunsystems, wie *PGRP-LB*, *PGRP-SC2* und *tollip*, sind Indikatoren einer reduzierten Immunantwort in den Ovarien von *C. floridanus*. Wie schon für den Mitteldarm der Tiere vorgeschlagen, könnte dies möglicherweise sowohl zur Toleranz von *Blochmannia* als auch zur vertikalen Übertragung der Endosymbionten beitragen.

Die vorliegende Doktorarbeit erweitert das Wissen über das Immunrepertoire von *C. floridanus* und es konnten vielversprechende Kandidaten für weitere funktionelle Analysen von möglichen Immunfaktoren identifiziert werden. Darüber hinaus konnten weitere Hinweise auf die Bedeutung von Immunfaktoren der Ameisen bei der Toleranz gegenüber den symbiontischen Bakterien gefunden und auf die Ovarien der Tiere ausgeweitet werden.

1 Introduction

1.1 The nature of bacterial endosymbioses in insects

Studying symbiotic interactions between microbes and their insect hosts provides insights into how bacteria infect insects, avoid the immune response, and even manipulate their hosts' physiology. With respect to the management of insect-related problems the increasing knowledge about such interactions and the development of methods to manipulate insect microbiota may allow (i) the birth control of agricultural insect pests; (ii) the control of insect vectors transmitting pathogenic diseases to humans, animals and plants; (iii) the protection of beneficial insects from parasites and pathogens or environmental stresses.

The nature of a symbiosis is defined as a close and constant living together of organisms belonging to different species. These associations may vary from mutualism to parasitism (De Bary, 1879). Depending on where the symbionts reside, they are referred to as endosymbionts when residing inside the host's body, whereas ectosymbionts inhabit the outer surface of their hosts. Symbioses at least partially facilitated diversification and the evolutionary success of insects (Buchner, 1965), which are reflected in their species richness and abundance as well as in the variety of inhabited ecological niches. It has been shown that bacterial symbionts contribute to their hosts' fitness in various ways (Feldhaar, 2011), for example via providing essential nutrients otherwise lacking in the insects' unbalanced diet (Akman et al., 2002; Douglas, 1998; Nogge, 1981; Zientz et al., 2004) or by facilitating the digestion or detoxification of food (Cardoza et al., 2006). Furthermore symbionts have been shown to provide thermal tolerance (Dunbar et al., 2007) as well as enhanced resistance to parasitoids and pathogens (Currie et al., 2003; Hamilton and Perlman, 2013; Hedges et al., 2008; Kaltenpoth, 2009; Oliver et al., 2010; Oliver et al., 2005) to their hosts. The host associated bacteria are traditionally classified into so-called primary and secondary endosymbionts.

Primary endosymbionts have obligate associations with their insect hosts that developed over long evolutionary timescales of approximately 30-250 million years (Baumann, 2005). Due to the obligate nature of the association, primary endosymbionts are transmitted strictly vertically via the germ line and their localisation within the insect hosts is restricted to ovaries as well as to specialised cells, the so-called bacteriocytes, that often constitute larger organ-like structures, the bacteriomes, which can be located adjacent to the ovaries (Braendle et al., 2003; Buchner, 1965). As a consequence of their intracellular life style, recombination with free-living bacteria is prevented and a high rate of genome evolution can be observed in primary endosymbionts. The accumulation of deleterious mutations, the drastic reduction of the genome size due to gene loss and a high adenine-thymine content are characteristics for most of the small genomes of primary endosymbionts (Klasson and Andersson, 2010; McCutcheon and Moran, 2012; Wernegreen, 2012). During coevolution of the hosts' and their symbiont's genomes, genes contributing to symbiosis, e.g. by complementing metabolic pathways, are

often retained within the genomes of endosymbionts, while genes not required in the stable intracellular environment are lost (Gil et al., 2004; Moya et al., 2008). Consequentially, primary endosymbionts have lost the ability to multiply independently from their hosts (Feldhaar and Gross, 2009) and are not culturable *in vitro* down to the present day (Kikuchi, 2009).

In contrast, secondary endosymbionts are facultative symbionts from the host's perspective and their localisation is not restricted to bacteriocytes, but they may be found intra- or extracellularly in various host tissues as well as in the haemocoel (Dale and Moran, 2006). Secondary endosymbionts are transmitted vertically, but also horizontally (Oliver et al., 2010; Russell et al., 2003). Although these facultative endosymbionts provide various benefits to their hosts' fitness, for example by endowing the insects with high-temperature tolerance (Montllor et al., 2002), host plant specificity (Tsuchida et al., 2004), or increased resistance to parasitoid wasps (Oliver et al., 2005; Oliver et al., 2003), they are not required for host survival or fecundity. Therefore, secondary endosymbionts may not be carried by all individuals of a host species (Harris et al., 2010; Kikuchi, 2009).

1.2   Strategies for the vertical transmission of insect endosymbionts

Especially in the cases of obligate, primary symbioses it is essential for the hosts' fitness to transfer the endosymbionts to the next generation as well as to every individual of one species. Differences in insect host biology and various localisations of the endosymbionts within the host facilitate several and highly diverse vertical transmission routes of endosymbiotic bacteria to the host's next generation. The nature of cells involved in transmission of endosymbionts also depends on the general morphology and physiology of the female reproductive organs, the ovaries.

Ovaries are most often pairwise and consist of bundles of ovarioles. The differentiation of germ-line cells into oocytes and the formation of follicles within these polar ovarioles are the distinguishing characteristic of ovaries. Each ovariole is positioned within the insects' abdomen by three somatic tissues, the envelope surrounding the whole ovary (peritoneal sheath), one envelope surrounding each ovariole (ovariolar epithelial sheath) and the terminal filament anchoring ovarioles anteriorly (Büning, 2005; Koch et al., 1967; Middleton et al., 2006). An ovariole can be divided into two parts: the anterior germarium, in which germ cells differentiate, and the posterior part, the so-called vitellarium, in which the growth of the oocyte and the yolk deposition into the growing oocytes occur (Büning, 2005). However, it is the first part, the germarium, which determines the final type of ovariole, of which there are two, the panoistic and the meroistic type (Brandt, 1874; do Amaral and Machado-Santelli, 2009) (Figure 1).

In panoistic ovarioles each germ cell can become an oocyte and therefore undergoes the prophase of meiosis within the germarium before entering the vitellarium. In meroistic ovarioles germ cells undergo a series of synchronised mitotic events resulting in clusters of mitotic siblings, which are termed cystocytes and remain connected to each other via intercellular bridges. Meiotic differentiation of cystocytes is initiated during multiplication and leads to the creation of one or several oocytes and nurse cells, also called trophocytes (Büning, 2005). Meroistic ovarioles can be classified into two categories: telotrophic and polytrophic (Gross, 1903). In telotrophic-meroistic ovarioles all prospective nurse cells (trophocytes) are kept in the germarium, which is turned into an anterior tropharium during the reproductive phases. Each developing oocyte remains connected to the tropharium via the nutritive cord. In contrast, in polytrophic-meroistic ovarioles the descendants of the germ cells, the cystocytes, differentiate into nurse cells and a single oocyte, which are kept together building a structural and functional cluster, the so-called egg-chamber (follicle). Afterwards, the whole egg-chamber enters the vitellarium and undergoes the processes of oocyte growth and vitellogenesis while being shifted towards the posterior end of the ovariole when younger egg-chambers are formed. During these processes nurse cells provide the oocytes with maternal factors. The whole germ cell cluster is enveloped by a single layer of follicle cells which have somatic origin and support the egg shell formation at the end of vitellogenic growth (Büning, 2005; Simiczyjew, 2003).
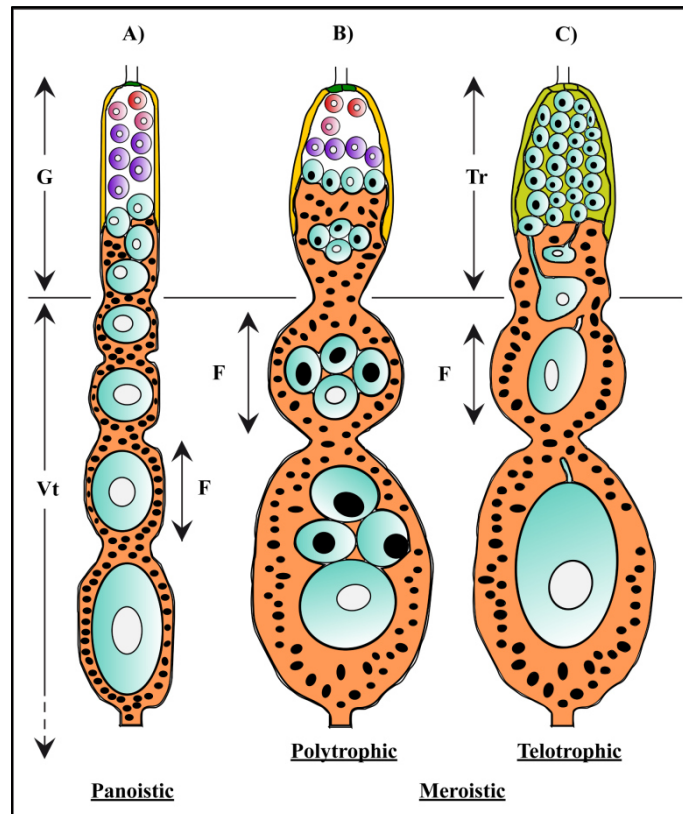
Figure 1 Overview over the three main types of ovarioles in insects. **(A)** In panoistic ovarioles, cap cells (green) interact with the germ-line stem cells maintaining them as stem cells (red). The panoistic ovariole type is generated by stem cells undergoing differential mitosis thereby producing cystoblasts (rose), which undergo further mitoses producing clusters of cystocytes (purple) in the germarium (G). All cystocytes can become oocytes (turquoise, white nuclei) which further grow in the vitellarium (Vt). In meroistic ovarioles, the differentiation into nurse cells (turquoise, black nuclei) and oocytes can be observed. This differentiation takes place in the germarium, which is sheathed by somatic inner sheath cells (yellow). **(B)** In polytrophic-meroistic ovarioles, only one cystocyte out of each cluster develops into an oocyte. The other cystocytes become nurse cells. Each cluster will be surrounded by a monolayer of follicle cells (brown). Such a developing unit is called the follicle (F), consisting of the oocyte-/ egg-chamber and its own nurse cell chamber. **(C)** In telotrophic-meroistic ovarioles, the apical cells comprise inner sheath cells and interstitial cells (light green) surrounding the anterior region, here called the terminal tropharium (Tr). As in panoistic ovarioles, each follicle consists of the oocyte-chamber only. Additionally, oocytes remain attached to the nurse cells (trophocytes) in the tropharium via the nutritive cord. (Figure adapted from (Büning, 2005)).

Follicle cells and the two types of germ cells, nurse cells and oocytes, have been shown to harbour endosymbionts during their transmission. However, the cell types infected with the endosymbionts and the point in time of transmission into those cells can differ drastically between different oviparous and viviparous insects. While the symbiotic bacteria have to be transmitted into the oocytes or eggs before oviposition in oviparous insects, the endosymbionts can be transferred even at later time points to the developing embryo inside viviparous insects.

The viviparous tsetse fly *Glossina morsitans* (Diptera) harbours a primary and several secondary bacterial endosymbionts. The obligate mutualistic endosymbionts *Wigglesworthia glossinidia* reside in the cytosol of bacteriocytes, which form a bacteriome in the anterior gut, and provide their hosts with vitamins lacking from the vertebrate blood diet (Akman et al., 2002; Chen et al., 1999). *W. glossinidia* as well as the secondary endosymbiont *Sodalis glossinidius*, that is present in the midgut epithelium

cells, are transmitted to the *in utero* developing embryo through milk gland secretions. A further secondary endosymbiont belonging to the genus *Wolbachia* already infects trophocytes and developing oocytes of the tsetse fly (Balmand et al., 2013). In the parthenogenetic aphid *Acyrthosiphon pisum* (Hemiptera), the primary endosymbiont *Buchnera aphidicola* supplies its host with essential amino acids lacking from the plant sap diet. The endosymbionts are located in vacuoles within bacteriocytes, which form a bacteriome in the haemocoel (Baumann, 2005) and are also transmitted to the *in utero* developing embryo of the aphid. The transport of *B. aphidicola* from the bacteriocytes into the blastulae (early embryos) proceeds in a series of exo- and endocytotic processes via an extracellular stage of the bacterial symbiont (Koga et al., 2012).

Ovipary is the more common mode of reproduction amongst insects. In contrast to the transmission of the endosymbionts in viviparous insects, which occurs most often in late stages during development of eggs and embryos, the infection of oocytes during transovarial transmission in oviparous insects is established either early or late during oogenesis. For example, the endosymbiont *Schneideria nysicola* of seed bugs of the genus *Nysius* (Hemiptera) is generally localised intracellularly in a pair of large bacteriomes which are in close association with the gonads. In adult female bugs *S. nysicola* was also found in the so-called infection zone, which is localised in the transition zone from germarium to vitellarium, where the first egg-chamber forms within the telotrophic-meroistic ovarioles of the seed bugs. Within the "infection zone" the *Schneideria* endosymbionts preferentially localised in the ovarial bacteriocytes. Additionally, *S. nysicola* was observed in the "symbiont balls" at the anterior poles of each oocyte, suggesting that the endosymbionts enter the developing oocytes ensuring vertical transmission of the bacteria. (Matsuura et al., 2012a; Matsuura et al., 2012b). In *Nysius ericae* and *Nithecus jacobaeae* (both Hemiptera), ovariolar bacteriocytes are located within the "infection zone" in the basal parts of the ovary. These specialised cells, surrounding early previtellogenic oocytes, harbour large amounts of endosymbiotic bacteria, which are gradually released from the bacteriocytes. Penetrating the oocytes plasma membrane, bacteria are eventually taken up by the oocyte and consequentially surrounded by an additional host-derived membrane. Being dispersed throughout the whole ooplasm at first, the endosymbionts accumulate at the anterior pole in vitellogenic oocytes forming a "symbiotic ball" (Swiatoniowska et al., 2013). Another example of early transfer of endosymbionts into oocytes can be found in scale insects. Similar to other plant sap-sucking insects, scale insects harbour obligate endosymbionts complementing their diet which is lacking certain essential amino acids (Baumann, 2005). In scale insects, endosymbionts may invade undifferentiated germ cells, such as cystocytes, but also oocytes in later stages of development (Buchner, 1965; Szklarzewicz et al., 2013; Szklarzewicz et al., 2006). In *Marchalina hellenica* all germ cells, trophocytes, arrested and developing oocytes, are interconnected forming a cluster of 25 to 32 cells. The trophocytes' cytoplasm is rich in endosymbiotic bacteria which are reproducing via binary fission and are separated from the cytoplasm by a host-derived membrane. During oogenesis, the bacteria are progressively transported from trophocytes into the developing oocyte (Szklarzewicz et al., 2013).

In contrast to very early and early infection of germ cells described above, several cases show the infection of oocytes through nurse cells in later stages of oogenesis or even early embryos. Various species of the parasitic chalcid wasps of the genus *Aphytis* (Hymenoptera) harbour endosymbionts of the genus *Wolbachia*, which are known to be responsible for several reproductive disorders. In the uniparental species, nurse cells of *A. yanonensis* are packed with *Wolbachia*. Maternal substances as well as *Wolbachia* are transferred into the young oocytes through cytoplasmic bridges (Zchori-Fein et al., 1998). Recently, a mutualistic bacterium in *Cardiocondyla obscurior* (Hymenoptera) was described. Queens, workers and wingless males of this invasive ant species harbour the γ-proteobacteria *Westeberhardia cardiocondylae* intracellularly in bacteriomes, which are connected to the gut. However, no such gut bacteriomes were detected in adult queens. Instead, *W. cardiocondylae* is highly abundant in ovaries of pupal and adult queens. There the endosymbionts are mostly located in the nurse cells and are transmitted into late-stage oocytes during nurse-cell depletion (Klein et al., 2016).

Yet another route of transmission was described in the leafhopper *Macrosteles laevis* (Hemiptera). Three types of bacterial endosymbionts are located intracellularly in bacteriocytes forming two large bacteriomes in close contact to the ovary. In the stage of early vitellogenesis, these endosymbionts leave the bacteriocytes and begin to invade the ovarioles. The symbiotic bacteria enter the follicle cells and temporarily reside in their cytoplasm before leaving the follicular cells and entering the perivitelline space, the space between the follicular epithelium and the cell membrane of the oocyte. There, the bacteria are located in a structure termed "symbiotic ball" and they invade the oocytes at their posterior pole only after the end of oocyte growth (Kobialka et al., 2015).

1.3   The insect immune system

Securing the transmission of endosymbionts does not only comprise the mere mechanisms of transmission. In fact, in each generation of the insect hosts, the endosymbionts have to establish chronic infections without being eliminated by their hosts' immune system. Since endosymbionts are often closely related to pathogenic bacteria (Moran et al., 2008), recognition and endosymbiont clearing by the insect immune system would be expected. However, host measures as well as specialised bacterial adaptations allow the endosymbionts to reside within the hosts in a controlled manner. Therefore, the host immune system plays a key role in establishing a stable infection of the endosymbionts within their insect hosts.

The insect immune system lacks the classical adaptive immunity known from vertebrates and therefore relies on the innate immune response only (Gross et al., 2009; Lemaitre and Hoffmann, 2007; Vallet-Gely et al., 2008). To date the innate immune response of *Drosophila melanogaster* is the most extensively studied and best-understood innate immune system of an insect. The innate immune response is generally characterised by a cellular and a humoral branch, both activated upon the recognition of microbe-associated molecular patterns (MAMPs). "Constitutive" defence mechanisms comprised of cellular actions, such as melanisation, phagocytosis, and the production of reactive oxygen species, are considered as a first line of defence, which is complemented by the humoral immune response. The latter mainly leads to the production of antimicrobial peptides (AMPs) at later time points of infection (Bulet et al., 2004; Kounatidis and Ligoxygakis, 2012; Lemaitre and Hoffmann, 2007; Strand, 2008). The main players in the cellular immune response are haemocytes, which are considered as the invertebrate equivalents to vertebrate blood cells. In *Drosophila*, three populations of functional haemocytes are produced from haemocyte precursors (pro-haemocytes) in a process called haematopoiesis which takes place during embryogenesis, larval development and throughout metamorphosis into adult animals (Kounatidis and Ligoxygakis, 2012; Vlisidou and Wood, 2015). Plasmatocytes are professional phagocytes, removing pathogens and apoptotic bodies via phagocytosis and comprise 90-95% of all mature larval haemocytes. Approximately 5% of haemocytes are crystal cells which are required for melanisation via the expression of components of the phenoloxidase (PO) cascade leading to melanin formation (Kounatidis and Ligoxygakis, 2012; Strand, 2008). Furthermore, a third population of haemocytes, the so-called lamellocytes, differentiates from the pro-haemocytes in response to specific immune challenges such as parasitic wasp infection. Lamellocytes encapsulate large foreign particles which cannot be phagocytosed (Kounatidis and Ligoxygakis, 2012; Lanot et al., 2001).

Humoral effectors, which comprise AMPs as well as factors involved in melanisation and clotting processes, are mainly produced in the insect fat body but also in haemocytes. As the cellular and the humoral immune response are highly interactive and influence each other, the distinction between the two branches is not quite strict (Feldhaar and Gross, 2008; Kounatidis and Ligoxygakis, 2012). The

first hallmark of either branch of the immune response in *Drosophila* is the recognition of MAMPs followed by the activation of downstream signalling cascades via different pathways: (i) the Toll pathway, (ii) the immune deficiency (IMD) pathway, (iii) the JAK/STAT (janus kinase/signal transduction and activator of transcription) and (iv) the JNK (c-jun N-terminal kinase) pathway. The main result of these pathways is the transcriptional regulation of the production of antimicrobial effectors via NF-κB-like transcription factors like Relish or Dorsal (Kounatidis and Ligoxygakis, 2012; Lemaitre and Hoffmann, 2007).

1.3.1 Microbial recognition

The recognition of "non-self" via MAMPs is facilitated by several pattern recognition receptors (PRRs) which mainly recognise and bind bacterial peptidoglycan (PGN) and fungal glucans. Peptidoglycan consists of a polymer of alternating *N*-acetylglucosamine and N-acetylmuramic acid residues which are cross-linked via peptide bridges. The latter are the characteristic distinguishing two types of PGN. Whereas Gram-negative bacteria generally contain a diaminopimelic acid (DAP)-type peptidoglycan, the PGN of almost all Gram-positive bacteria features cross-linking via lysine (Kurata, 2014). Peptidoglycan is recognised by PRRs called peptidoglycan recognition proteins (PGRPs) which can be categorised into two subgroups according to either structure or function. Short *pgrp* genes encode extracellular proteins (PGRP-S) while long *pgrp* genes code for intracellular or membrane spanning proteins (PGRP-L). Several long and short PGRPs (PGRP-LB, PGRP-SB1, PGRP-SB2, PGRP-SC1(a/b) and PGRP-SC2) contain a domain homologous to T7 lysozyme and therefore display a zinc-dependent amidase activity hydrolysing DAP-type PGN. As the degraded PGN shows no immunostimulatory activity, it is suggested that the scavenger activity of such PGRPs prevents an excessive activation of the immune response via modulation of the IMD pathway (Bischoff et al., 2006; Kurata, 2014; Paredes et al., 2011; Royet et al., 2011; Zaidman-Remy et al., 2006; Zaidman-Remy et al., 2011). Another group of PGRPs cannot hydrolyse PGN due to the lack of a cysteine residue critical for the amidase domain activity. Upon binding of such sensor PGRPs to either lysine-type PGN (PGRP-SA and -SD) or to DAP-type PGN (PGRP-LC and -LE) the Toll and the IMD pathway are activated (Kurata, 2014; Lemaitre and Hoffmann, 2007). Interestingly, PGRP-LE is a constitutively expressed haemolymph protein and recognises extracellular "non-self" particles. Additionally, the *Drosophila* PGRP-LE acts as a co-receptor of the surface-bound PGRP-LC and as an intracellular receptor in immune reactive cells (Kaneko et al., 2006; Kurata, 2010; Takehana et al., 2002; Takehana et al., 2004).

A further set of PRRs shows similarities to bacterial glucanases. In *Drosophila*, the so-called Gram-negative binding proteins (GNBPs) bind to bacterial lipopolysaccharides, lipoteichoic acids or fungal β-1,3-glucans initiating the activation of the Toll pathway (Gottar et al., 2006; Kurata, 2014). The *Drosophila* GNBP1 shows muramidase-like activity hydrolysing lysine-type PGN into short

muropeptide ligands for PGRP-SA and is required for PGRP-SA dependent recognition of Gram-positive bacteria (Wang et al., 2006).

1.3.2 The three main immune signalling pathways

The *Drosophila* genome encodes three NF-κB-like transcription factors, Relish, Dorsal and Dif, which bind to κB sites thereby activating AMP gene expression. These transcription factors play key roles in the AMP expression regulation in two of the major immune signalling pathways (Figure 2).
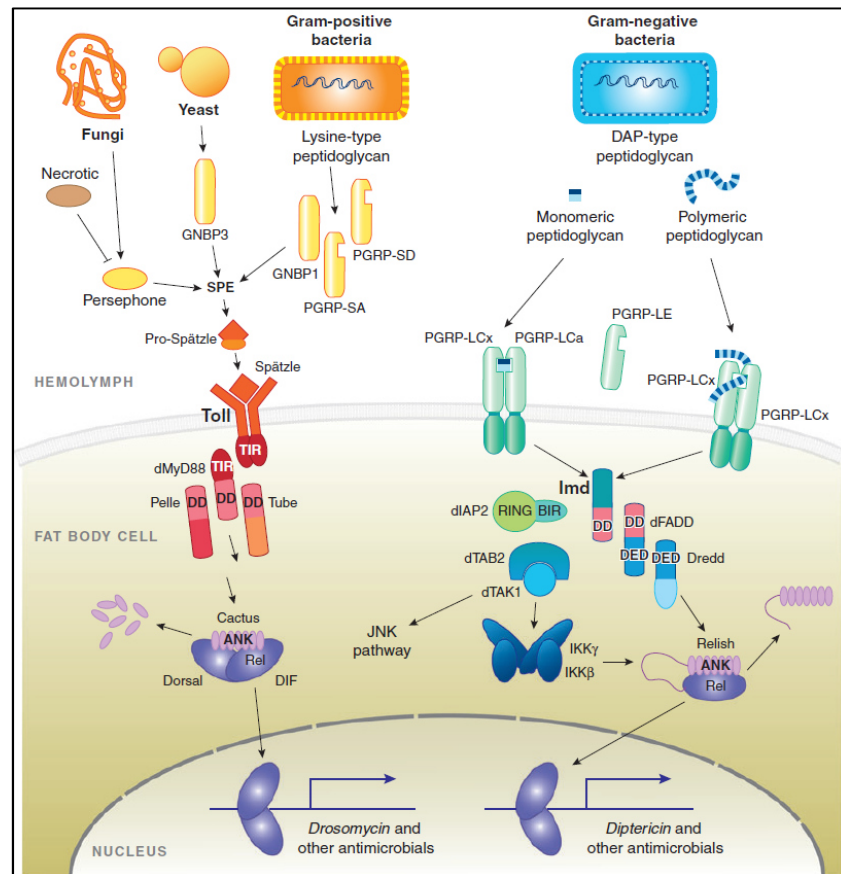
Figure 2 Model of Toll and IMD pathway activation. In *Drosophila*, two main signalling pathways regulate the expression of AMPs: the Toll pathway and the IMD pathway. The *Toll pathway* is mainly activated upon binding of secreted PRRs sensing Gram-positive bacteria (PGRP-SA, PGRP-SD, GNBP1), yeast Glucans (GNBP3) and entomopathogenic fungi (Persephone). A proteolytic cleavage cascade leads to cleavage and activation of Spätzle by the Spätzle processing enzyme (SPE). Upon binding of Spätzle to Toll, dimerization of the receptor is induced and three intracellular death domain- (DD-) containing proteins, MyD88, Tube, and Pelle, are recruited. The protein Cactus is phosphorylated and degraded by the proteasome. The Rel transcription factors Dif and Dorsal are released and move from the cytoplasm into the nucleus, where they regulate the expression of several antimicrobial effectors (Lemaitre and Hoffmann, 2007; Valanne et al., 2011). *IMD pathway*: Monomeric or polymeric DAP-type PGN of Gram-negative bacteria is recognised by membrane-bound PGRP-LC, which then recruits a signalling complex consisting of Imd, a death domain protein homologous to mammalian RIP1, the adaptor protein dFADD, and the caspase Dredd. Dredd becomes activated upon ubiquitination and cleaves Imd, which leads to recruitment and activation of the Tab2/Tak1 complex and further to the phosphorylation of the *Drosophila* IKK complex. Via phosphorylation of multiple sites, Relish is activated and its N-terminal Rel domain is translocated to the nucleus, where the transcription of target genes encoding for example AMPs is activated (Lemaitre and Hoffmann, 2007; Myllymaki et al., 2014). (TIR = Toll-IL1 receptor domain; DED = death-effector domain; ANK = ANKyrin repeats; Rel = Rel homology domain; RING = RING finger domain; BIR = Baculovirus IAP repeat). (Figure modified from (Lemaitre and Hoffmann, 2007); use of the Figure under the terms of the licence agreement with Annual Reviews, Inc provided by Copyright Clearance Center, Appendix XI)

The Toll pathway is required for haemocyte proliferation and regulates the antimicrobial immune response (Lemaitre and Hoffmann, 2007). It also plays a role in the establishment of the dorso-ventral axis during *Drosophila* embryogenesis (Qiu et al., 1998). Upon binding and recognition of "non-self" structures via PGRP-SA, PGRP-SD, GNBP1 and GNBP3, Spätzle is activated in a proteolytic cascade and binds to the membrane-bound Toll receptor. Dimerization of the receptor causes the recruitment of three intracellular death domain-containing proteins, MyD88, Tube, and Pelle. Further downstream, the protein Cactus is phosphorylated and degraded by the proteasome resulting in the release of the Rel transcription factors Dif and Dorsal, which then move from the cytoplasm into the nucleus, where they regulate the transcription of several antimicrobial effectors (Aggarwal and Silverman, 2008; Lemaitre and Hoffmann, 2007; Valanne et al., 2011).

The IMD pathway is mostly activated upon binding of monomeric or polymeric DAP-type PGN of Gram-negative bacteria to the membrane-bound PGRP-LC. A signalling complex consisting of Imd, a death domain protein homologous to mammalian RIP1, the adaptor protein dFADD and the caspase Dredd is then recruited. Dredd becomes activated via ubiquitination and cleaves Imd resulting in the recruitment and activation of the Tab2/Tak1 complex. The *Drosophila* IKK complex is phosphorylated and again induces phosphorylation of multiple sites of Relish which is thereby activated. Only the N-terminal Rel domain is translocated to the nucleus to activate the transcription of target genes encoding AMPs and other immune effectors (Kleino and Silverman, 2014; Lemaitre and Hoffmann, 2007; Myllymaki et al., 2014). A side effect of the activation of the IMD pathway is the Tak1 (transforming growth factor β-activated kinase 1) mediated activation of the JNK pathway which contributes to several developmental processes (Noselli and Agnes, 1999), wound healing (Rämet et al., 2002a) and cellular immune response in *Drosophila* (Park et al., 2004; Sluss et al., 1996). The JNK kinase is activated via Tak1 and initiates phosphorylation and nuclear translocation of the transcription factor AP-1 resulting in immune gene expression (Kounatidis and Ligoxygakis, 2012; Park et al., 2004).

The JAK/STAT pathway has been characterised primarily in *Drosophila*. The main cellular components comprise the receptor Domeless, the Janus kinase (JAK) Hopscotch, the transcription factor STAT and three Unpaired-related ligands (Upds) (Binari and Perrimon, 1994; Brown et al., 2001; Harrison et al., 1998; Hombria et al., 2005; Yan et al., 1996). The binding of Upd induces dimerization of the receptor Domeless which allows phosphorylation of Hopscotch. In the following, the STAT transcription factor is recruited and phosphorylated. Upon dimerization of STAT, the transcription factor is transported to the nucleus and regulates the expression of effector genes (Kingsolver et al., 2013). In *Drosophila*, it is known that the gene encoding the complement-like protein Tep2 (thioester-containing protein) and the *turandot* stress genes are regulated by the JAK/STAT pathway upon septic injury (Agaisse and Perrimon, 2004; Agaisse et al., 2003). Furthermore, the transcriptional regulation of these genes receives additional inputs from the IMD and

the MAPK (mitogen-activated protein kinase) pathways (Brun et al., 2006). The JAK/STAT pathway was originally identified through its role in embryogenesis, but it has been shown that the pathway responses to tissue damage during infection, various stress conditions during septic injuries and to bacterial as well as to *Drosophila* C virus infections (Agaisse and Perrimon, 2004; Brun et al., 2006; Dostert et al., 2005; Kingsolver et al., 2013; Lemaitre and Hoffmann, 2007).

1.3.3 Phagocytosis

The receptor mediated recognition and removal of cell fragments called apoptotic bodies and of bacterial or fungal infections is a crucial part of the first line of defence in insects. Professional phagocytes attach to the targeted particle which is internalised through cytoskeleton modification and vesicle trafficking within the phagocyte. The engulfed target is then degraded within phagosomes (Agaisse and Perrimon, 2004; Vlisidou and Wood, 2015). During development, the highly conserved *Drosophila* receptors Croquemort, NimC4/Simu and Draper recognise dead cells. The latter also recognises lipoteichoic acid from *Staphylococcus aureus*, which mediates the uptake of the bacterium (Hashimoto et al., 2009). Although the question of which bacterial or fungal components are recognised is still unanswered, several phagocytic receptors relevant to *Drosophila* immunity have been identified. In larvae and adults, members of the Nimrod protein family, Eater (Kocks et al., 2005) and Nimrod C1 (Kurucz et al., 2007), recognise Gram-positive and Gram-negative bacteria. Receptors of the scavenger receptor family Peste mediate phagocytosis of *Mycobacterium fortuitum* (Philips et al., 2005), whereas PGRP-LC mediates the uptake of *Escherichia coli* (Rämet et al., 2002b). Additionally, haemocyte-specific loss of the receptor Dscam (Down syndrome cell adhesion molecule) has been shown to cause impaired efficiency in the phagocytosis of *E. coli* (Watson et al., 2005). Finally, haemocytes express the so-called thioester-containing proteins (TEPs) which are evolutionary related to mammalian complement factors. The latter have been shown to mediate the opsonisation of pathogens, thereby increasing the phagocytosis of bacteria and fungi. Several TEPs can bind to *E. coli* and *S. aureus* promoting phagocytosis by *Drosophila S*2 cells (Stroschein-Stevenson et al., 2006). However, TEPs might act redundantly during infection, since flies lacking one or several *Tep* genes were not more susceptible to infection (Bou Aoun et al., 2011).

1.3.4 Melanisation

Haemolymph clotting and melanisation are activated immediately upon parasite or pathogen induced physical breakage of the insect cuticle. The main characteristic of melanisation, which is involved in wound healing, encapsulation and sequestration of microorganisms, is a blackening reaction due to synthesis of melanin pigments (Lemaitre and Hoffmann, 2007; Söderhäll and Cerenius, 1998). Whereas melanisation is performed solely by crystal cells in *Drosophila* larvae, the type of cell producing active phenoloxidase (PO) in adult animals remains unknown. Crystal cells produce the enzymatically inactive precursor prophenoloxidase (proPO). The *Drosophila* genome encodes three proPOs. Physical injury and the recognition of certain MAMPs, such as lipopolysaccharides (LPS),

PGN, and β-1,3-glucans, trigger the melanisation or proPO cascade (Kanost et al., 2004; Lee et al., 2004; Ochiai and Ashida, 2000). A serine protease (SP) called prophenoloxidase activating enzyme (PPAE) activates PO via cleavage of its precursor. The PPAE activity itself is also stimulated in a stepwise process involving serine proteases (Cerenius and Söderhäll, 2004). The tyrosine hydroxylase (TyrOH) hydroxylates the monophenol tyrosine. The resulting ortho-phenol DOPA (3,4-dihydroxyphenylalanine) is converted into dopamine, which is the substrate of the PO. Polymerisation of dopamine leads to the formation of melanin (De Gregorio et al., 2001). Additionally, it has been shown that several cytotoxic melanin precursors and other products like reactive oxygen species are produced. Such intermediates are directly toxic to various microbes (Cerenius and Söderhäll, 2004; Nappi and Vass, 1993; Volz et al., 2006). In order to prevent systemic melanisation, the activation of the melanisation cascade is regulated by SP inhibitors called serpins (De Gregorio et al., 2002).

1.3.5 The systemic immune response – fat body AMP production

Antimicrobial peptides (AMPs) are natural substances with antimicrobial functions and are synthesised by a broad spectrum of organisms including bacteria, plants and animals. The production and secretion of AMPs by the insect fat body and haemocytes display a hallmark of the systemic immune response in insects although AMPs generally appear to be of importance at later time points of immunity. It was shown that this branch of the immune response is required for killing of bacteria which survived the host's immediate defence mechanisms like phagocytosis or melanisation (Haine et al., 2008). Since the first purification and description of an insect AMP (cecropin) (Steiner et al., 1981), more than 150 insect AMPs have been identified. Most insect AMPs are small with a quite low molecular weight (< 10 kDa) and display cationic/ basic and hydrophobic properties. They show membrane-activity against bacteria and/or fungi and can also have activity against parasites and viruses. Due to certain structures or unique sequence motifs, AMPs are classified into four groups: (i) α-helical peptides (e.g. cecropin and moricin), (ii) cysteine-rich peptides (e.g. insect defensin and drosomycin), (iii) proline-rich peptides (e.g. apidaecin and drosocin), and (iv) glycine-rich proteins (e.g. attacin and gloverin). Based on these structural features and their general properties, two main mechanisms for interaction or integration of AMPs into target membranes are discussed: (i) the barrel stave model describing the formation of transmembrane channels/ pores due to AMP polymerisation with the hydrophobic surfaces of AMPs interacting with the lipid core of the membrane and the hydrophilic surfaces forming an aqueous pore; and (ii) the carpet model with the AMPs being in contact with the phospholipid head group on the outer leaflet of the membrane, which causes membrane permeation via disruption of the bilayer curvature (Reddy et al., 2004; Shai, 1999). The expression of AMPs is regulated by the Toll and IMD signalling pathways and most AMPs are synthesised as inactive precursors or proproteins, and active peptides are only generated upon proteolysis (Bulet and Stocklin, 2005; Yi et al., 2014).

So far, several AMPs from Hymenopteran species have been reported (Casteels-Josson et al., 1994; Choi et al., 2008; Evans et al., 2006; Ratzka et al., 2012a; Tian et al., 2010; Viljakainen and Pamilo, 2008). An bioinformatical approach also led to the prediction of the existence of tachystatin-like and crustin-like AMPs in seven ant genomes (Zhang and Zhu, 2012). Amongst the most often described AMPs in Hymenopterans are defensins and hymenoptaecins.

Defensins are generally short (~ 4 kDa) cationic peptides containing six conserved cysteine residues which form three intramolecular disulphide bridges. The structure of such defensins comprises an N-terminal loop, an α-helix, and an antiparallel β-sheet. Two intramolecular disulphide bonds link the α-helix and the antiparallel β-sheet, which leads to the formation of a "loop-helix-beta-sheet" structure (Cornet et al., 1995). Most insect defensins are active mainly against Gram-positive bacteria, although it has been reported that some insect defensins are also active against Gram-negative bacteria and fungi (Chapuisat et al., 2007; Yi et al., 2014). In the ant *Camponotus floridanus*, which was the model system of this work, two different sequences encoding defensin-like AMPs, which are homologs of the *Apis mellifera* defensin-1 and defensin-2, were identified (Ratzka et al., 2012a). Both genes encode prepropeptides each comprising a signal peptide, a propeptide and the following mature peptide. It was shown that especially the expression of *defensin-1* is significantly upregulated in fat body and midgut tissue of *C. floridanus* after immune challenge with Gram-positive and Gram-negative bacteria (Ratzka, 2012).

Hymenoptaecins are glycine-rich AMPs with activity against Gram-positive and Gram-negative bacteria. So far, they have only been found in Hymenopteran species and it has been shown that the *A. mellifera* hymenoptaecin causes lethal permeabilisation of the outer and inner membrane of *E. coli* (Casteels et al., 1993). The *hymenoptaecin* gene in *A. mellifera* encodes a propeptide with a signal and a pro-sequence. Upon processing of the precursor, the active hymenoptaecin peptide, which consists of 93 amino acids (aa) (10 kDa), is released (Casteels-Josson et al., 1994; Casteels et al., 1993). The *C. floridanus hymenoptaecin* gene codes for a multipeptide precursor (Figure 3) consisting of a signal peptide (19 aa), a propeptide (26 aa), and a putative hymenoptaecin-like domain peptide (108 aa) followed by six hymenoptaecin repeat domain peptides (97 aa each). The seven putative hymenoptaecin peptides are separated from each other by a spacer sequence EAEP (glutamic acid-alanine-glutamic acid-proline) and a putative processing site consisting of two arginine residues (RR) or a combination of lysine and arginine (KR) (Ratzka et al., 2012a). The hymenoptaecin multipeptide precursor shows high structural similarities to the precursor molecule of the *A. mellifera* AMP apidaecin. It was proposed that the apidaecin multipeptide precursor is processed in a manner similar to the maturation of the yeast α-factor (Casteels-Josson et al., 1993). The same mechanism involving a *KEX* encoded endoprotease and a carboxypeptidase (Dmochowska et al., 1987; Fuller et al., 1989) was suggested for the maturation of the *C. floridanus* hymenoptaecin peptides since homologs of the respective yeast KEX proteins were found in the ant's genome (Bonasio et al., 2010; Ratzka et al.,

2012a). The expression of the *hymenoptaecin* gene was highly upregulated after bacterial infection in *C. floridanus* (Ratzka et al., 2011) and processing of the multipeptide precursor would therefore result in an amplification of the immune response by releasing seven putative hymenoptaecin peptides (Ratzka et al., 2012a).
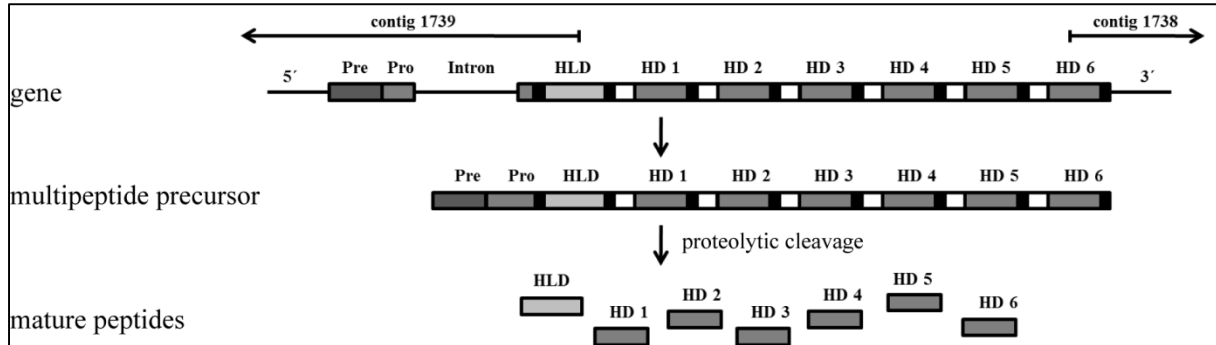
Figure 3 Schematic structure of the *hymenoptaecin* gene. The *hymenoptaecin* gene encodes a multipeptide precursor protein which consists of a signal-sequence (Pre) and a pro-sequence (Pro) followed by a hymenoptaecin-like domain (HLD) and six repeated hymenoptaecin domains (HD 1–6). The hymenoptaecin domains are flanked by the two putative processing sites "EAEP" (white boxes) and "RR" (black boxes). It was suggested that proteolytic cleavage of the multipeptide precursor molecule would result in the release of seven mature hymenoptaecin peptides with putative antimicrobial function (Ratzka et al., 2012a).

1.3.6 Social Immunity

One of the main reasons for the dominance of some insect species in many habitats is their social lifestyle. Social insects - ants, termites, and some wasps and bees – benefit from division of labour, cooperation in brood care, foraging, or anti-predator defences (Wilson, 1971). However, the social lifestyle comes with several potential drawbacks. Infectious diseases may spread faster amongst thousands of genetically closely related individuals and colonies might also face specific problems concerning nest and brood hygiene. Several behavioural and organisational adaptations in order to overcome such problems are referred to as social immunity (Cremer et al., 2007; Le Conte et al., 2011; Otti et al., 2014).

By way of example, it was shown that termites and ants not only avoid known areas rich in parasites but also stop devouring animals from their own colony (intraspecific cannibalism), when these individuals are infected with parasites (Cremer et al., 2007; Epsky and Capinera, 1988; Kramm et al., 1982; Zhou et al., 2002). Furthermore, several bees, termites and ants collect antimicrobial substances from the environment or use self-produced chemicals and mix them into the nest material for disinfection (Christe et al., 2003; Gilliam et al., 1988; Ortius-Lechner et al., 2000; Rosengaus et al., 1998), and infected brood is efficiently detected and removed from the nests in order to prevent spreading of parasitic infections (Cremer et al., 2007). Additionally, passive social immunisation was suggested as protective mechanism against bacterial infection. Workers in ant colonies of *Camponotus pennsylvanicus* which were abdominally injected with bacteria or LPS displayed

increased trophallactic behaviour. The regurgitate droplets showed increased antimicrobial activity and it was suggested that resistance at a colony level may be achieved via the trophallactic spreading of antimicrobial-mediated disease resistance (Hamilton et al., 2011). Also, active social immunisation was observed in *Lasius* ant colonies (Konrad et al., 2012). Upon fungal infection, an increased level of allogrooming behaviour between fungus-exposed and healthy nestmates was shown to trigger low-level infection in healthy animals. These infections through fungal transfer are non-lethal but induce an immune response increasing the ability to inhibit fungal growth on a colony level. However, it appears that the active immunisation of the colony is infection-specific. All in all, it was suggested that active social immunisation is adaptive and leads to faster and more effective elimination of diseases than passive immunisation (Konrad et al., 2012).

1.4   Endosymbiont tolerance and regulation via the insect host immune system

It is quite obvious that insects harbouring endosymbionts require special adaptations regarding a possible immune response against their symbionts. At a glance, it seems quite surprising that insects are capable of fighting a pathogenic infection while maintaining a stable infection with their symbiont. In addition, bacterial endosymbionts are often closely related to pathogenic bacteria (Moran et al., 2008), which are generally recognised and fought by the immune system. However, several studies have shown adaptations in the symbiotic bacteria as well as in their hosts promoting symbiont tolerance and regulation within the insects (Ratzka et al., 2012b).

One of the first questions arising is whether endosymbiotic bacteria are recognised by their host's immune system or not. Moreover, do endosymbionts actively manipulate the host's immune system in favour of their survival. For example it was shown that infection of *D. melanogaster* with two *Spiroplasma* strains, the MSRO (melanogaster sex ratio organism) and the NSRO (nebulosa sex ratio organism), did not induce the expression of several immune genes, e.g. those coding for AMPs (Herren and Lemaitre, 2011; Hurst et al., 2003; Hutchence et al., 2011). The male-killing and therefore sex ratio distorting *Spiroplasma* endosymbionts colonise mainly the haemolymph as well as several tissues within different *Drosophila* species (Montenegro et al., 2005; Sakaguchi and Poulson, 1961). It may be that the cell-wall-less *Spiroplasma* (Whitcomb and Williamson, 1975) does not trigger an immune response due to the lack of MAMPs such as PGN, which usually activate immune signalling pathways (Lemaitre and Hoffmann, 2007). Additionally, resistance of *Spiroplasma* to the *D. melanogaster* immune system seems to be strain specific. Whereas the NSRO *Spiroplasma* strain was killed in *D. melanogaster* mutants with a constitutively active Toll signalling pathway (Hurst et al., 2003), the MSRO strain was not susceptible to Toll and IMD mediated immune responses (Herren and Lemaitre, 2011). It was also shown that a male-killing NSRO *Spiroplasma* strain suppresses immune gene expression in immune challenged insects, while a non-male-killing NSRO-A strain is not capable of such immune suppression and cannot withstand immune defences in *D. melanogaster* (Anbutsu and Fukatsu, 2010).

The negative regulation of the immune signalling prevents an overactive immune response which might be harmful to the animals. It may also modulate the immune response in favour of endosymbiont tolerance. For example, it was shown that especially PGRP-LB and -SC1/2 are expressed in the *Drosophila* gut supposedly modulating the immune response towards commensal gut bacteria. These PGRPs contain amidase domains which catalyse the cleavage of peptidoglycan into non-immunostimulating particles. This negative immune regulation is suggested to allow the stable infection with commensal gut bacteria (Bischoff et al., 2006; Zaidman-Remy et al., 2006).

Interestingly, bacteriocytes or bacteriomes harbouring endosymbionts can be a "safe harbour" for symbiotic bacteria. In the maize weevil *Sitophilus zeamais*, two putative negative immune regulators, tollip and wPGRP-1, a homolog to PGRP-LB, are constitutively overexpressed in the bacteriome

harbouring the *Sodalis pierantonius* str. SZPE (*Sitophilus zeamais* primary endosymbiont) (Anselme et al., 2008; Heddi et al., 2005; Login et al., 2011). Additionally, it was shown that the amidase wPGRP-1 was not only highly expressed in the bacteriome, but wPGRP-1 expression was also strongly induced in the symbiotic nymphal phase of *S. zeamais*. In this phase, the symbionts are released from the bacteriocytes and it is believed that wPGRP-1 activity prevents immune defences against the transitory extracellularly symbionts by removing bacterial immunostimulants (Anselme et al., 2006). However, it was shown that isolated *S. pierantonius* str. SZPE, when injected into the adult weevil haemolymph, was recognised inducing a strong induction of AMP gene expression. In contrast, only one AMP gene encoding coleoptericin-A (ColA) was strongly expressed in the symbiont bearing bacteriome (Login et al., 2011). Immunohistochemical analyses revealed the expression of ColA in all tissues harbouring *S. pierantonius* str. SZPE including the thin layer of follicle cells surrounding oocytes, which indicates a role of ColA in confining the endosymbiont within the bacteriome and oocytes. It was also demonstrated that the AMP interacts with the outer membrane proteins OmpA and OmpC of *S. pierantonius* str. SZPE and with the intracellular bacterial chaperonin GroEL (Login et al., 2011). The ColA-GroEL interaction may interfere with bacterial cell division as it has been shown for temperature-sensitive *groEL* mutations in *E. coli* triggering cell elongation. The chaperonin co-localises with the FtsZ-ring, the bacterial cell division machinery. In the absence of GroEL the FtsZ-ring is instable, which is why it was suggested that the chaperonin is either necessary for recruitment of other cell division proteins or a general stabiliser of the FtsZ-ring (Ogino et al., 2004). Furthermore, the folding of an important component of the FtsZ-ring, FtsE, is strictly dependent on GroEL and impaired folding causes the filamentous *E. coli* phenotype (Fujiwara and Taguchi, 2007). Interestingly, *S. pierantonius* str. SZPE cells inside the bacteriocytes are highly polyploid and very long (up to 200 μm). Bacterial length and polyploidy correlate directly which suggests a disturbance in cell division. It was therefore speculated that ColA interacts with SZPE GroEL leading to misfolding of FtsE and impaired cell division. Thus, the localisation of the elongated endosymbionts is restricted to the bacteriocytes, since the bacteria cannot escape these cells due to their size (Login et al., 2011; Login and Heddi, 2013). Recently, it has been demonstrated that immune response mechanisms of *S. zeamais* are also involved in the elimination of the *S. pierantonius* str. SZPE. Vigneron *et al.* (2014) observed the rapid decrease of the endosymbiont population in the gut between day six and day 14 of adulthood only shortly after a massive multiplication of the bacterial symbionts in young adult weevils, while the ovarian symbiont population remained stable (Vigneron et al., 2014). It was shown that apoptosis is induced in bacteriocytes of symbiont bearing animals but not in aposymbiotic weevils, indicating the activation of apoptosis specifically addressing symbiont elimination. Additionally, autophagy mechanisms, which maintain cell homeostasis by elimination and recycling of defective organelles, were observed in symbiotic weevils (Vigneron et al., 2014). It is known that *S. pierantonius* str. SZPE provides its host with nutrients underrepresented in their natural diet, especially with the aromatic amino acids phenylalanine and tyrosine (Oakeson et al., 2014). It was

demonstrated that symbiont-dependent complementation of the diet with components other than carbohydrates is necessary especially in this critical phase of early adulthood. Aposymbiotic animals died more rapidly particularly under high temperatures and dry conditions pointing towards a relation to cuticle defects (Vigneron et al., 2014). The amino acids tyrosine and phenylalanine are substrates in the synthesis of DOPA the precursor molecule for melanisation and cuticle formation. High performance liquid chromatography (HPLC) analyses revealed a strong increase in the level of DOPA in symbiotic animals after day five. Complementation of the diet of aposymbiotic weevils with tyrosine was sufficient to achieve the same phenotype. These results suggest a very fine tuned regulation of endosymbiont populations in *S. zeamais* according to their needs. Tyrosine produced by *S. pierantonius* str. SZPE is obtained via digestion of the endosymbionts within weevils after day five of adulthood facilitating proper and rapid melanisation and cuticle formation in a life stage of high vulnerability to environmental conditions (Vigneron et al., 2014).

Finally, it was shown that immune gene regulation supports the transmission of endosymbionts in the tsetse fly *Glossina morsitans morsitans*. As aforementioned, the primary endosymbiont *W. glossinidia* is transmitted to the offspring via milk gland secretions during intrauterine development of the embryos (Attardo et al., 2008; Balmand et al., 2013). Additionally, PGRP-LB was shown to be a major component of the milk and RNAi mediated knockdown of PGRP-LB expression resulted in decreased *Wigglesworthia* number in the bacteriome and milk. Consequently, it was assumed that the amidase PGRP-LB cleaves *Wigglesworthia* PGN into non-immunostimulants in wild-type animals. Therefore, the reduced expression of immune-related genes such as AMP genes contributes to symbiont survival and transmission via the mother's milk (Wang and Aksoy, 2012). Overall, mutual or reciprocal adaptations with respect to the host's innate immune response contribute to the transmission and manifestation of stable infections with commensal and symbiotic bacteria in insects.

1.5 The symbiosis of *Camponotus floridanus* and *Blochmannia floridanus*

Already in the late 19th century, Friedrich Blochmann described the first bacteriocyte-associated symbioses in carpenter ants of the genus *Camponotus* (Blochmann, 1882). This ant genus belongs to the most successful ant genera due to its species richness with more than 1,000 described species and a worldwide distribution. Most *Camponotus* species are omnivorous and have a monogynous lifestyle with only one queen per colony but thousands of workers displaying a distinct size polymorphism (Bolton, 1995). Evolutionary success of the genus *Camponotus* is represented by species richness and habitat variety and might be supported by their bacterial symbioses with γ-Proteobacteria of the genus *Blochmannia*. So far, endosymbionts of the genus *Blochmannia* have been found in every *Camponotus* species investigated (Blochmann, 1882; Buchner, 1918; de Souza et al., 2009; Degnan et al., 2004; Feldhaar et al., 2007; Sameshima et al., 1999; Sauer et al., 2000). The Gram-negative, rod-shaped symbiotic bacteria reside mostly in the cytoplasm of the bacteriocytes which are intercalated in between midgut tissue cells. In female ants, *Blochmannia* can also be found in matured oocytes (Sauer et al., 2002; Schröder et al., 1996). The bacterial endosymbionts are transmitted vertically emphasising strict coevolution of the host and symbiont species. Therefore, *Blochmannia* species are classified according to their host species (Degnan et al., 2004). The endosymbiont of *Camponotus floridanus* is named *Blochmannia floridanus*.

B. floridanus and other *Blochmannia* species form a taxon phylogenetically closely related to other known symbiont taxa, e.g. the symbionts of tsetse flies or aphids. They are also close relatives to pathogenic members of the *Enterobacteriaceae*, for example *E. coli* and *Salmonella typhi* (Gil et al., 2003; Sauer et al., 2000). The genome of *B. floridanus* consists of a small (about 705 base pairs) circular chromosome which shows low GC-content (27.38%) and gene loss, typical for symbionts. Interestingly, the genome carries none of the genes encoding the replication initiation factors DnaA, PriA or RecA. It was suggested that replication of the cytosolic endosymbionts requires either alternative bacterial replication factors such as HplA or the direct control by host cell factors. In contrast, *Buchnera*, a symbiont enclosed with host membranes within the bacteriocytes, retains the *dnaA* gene for replication initiation (Gil et al., 2003). The genome of *B. floridanus* also lacks genes coding for proteins involved in biosynthetic pathways for several non-essential amino acids. However, the endosymbiont has retained biosynthetic pathways for almost all essential amino acids, and the genome codes for a complete urease gene cluster as well as for all enzymes involved in sulphur recycling. It was therefore proposed that *Blochmannia* contributes to host nitrogen and sulphur recycling (Gil et al., 2003; Zientz et al., 2004). The endosymbiont's role in nutritional upgrading of their hosts' diet was further analysed (Feldhaar et al., 2007; Zientz et al., 2006). It was demonstrated that aposymbiotic workers exclusively fed with a diet lacking essential amino acids raised significantly less brood than either groups of workers bearing the endosymbiont or groups of aposymbiotic workers fed with a diet also providing essential amino acids. These experiments showed that *Blochmannia* provide essential amino acids to their hosts (Feldhaar et al., 2007).

During the ants' metamorphosis, the distribution of bacteriocytes within the midgut and the number of *Blochmannia* are highly dynamic. Both, the number of bacterial endosymbionts and the number of bacteriocytes, peak in the late pupal stage, when the entire midgut is transformed into a symbiotic organ (Stoll et al., 2010). Since the number of bacteriocytes and symbionts decreased constantly in adult animals, it was suggested that *Blochmannia* may especially contribute to their hosts' development during metamorphosis. Accordingly, particularly in the pupal stage an upregulation of expression of several putative symbiotic genes (e.g. urease associated genes, genes involved in amino acid synthesis) was observed via transcriptional profiling of *B. floridanus* genes. Taken together these results suggest a contribution of the symbiont to its host via nutritional upgrading in a life phase during which no external food is ingested (Stoll et al., 2010; Stoll et al., 2009).

Recent studies shed some light on the involvement of the *C. floridanus* immune system in the regulation of its endosymbiont. It has been demonstrated that septic injury with pathogenic bacteria but also with the endosymbiont itself induces the expression of several immune-related genes. Such genes are involved in recognition (e.g. *GNBP*), immune signalling (e.g. *stubble*) and antimicrobial activity (e.g. *defensin* and *hymenoptaecin*) (Ratzka et al., 2011). The fact that the endosymbiont is still recognised by its host's immune system and induces a proper immune response led to the question whether there are special adaptations promoting endosymbiont tolerance in the midgut. Therefore, the expression of several candidate genes with putative immune function in the midgut tissue when compared to the residual body was observed over the stages of *C. floridanus* ontogeny (Ratzka et al., 2013a). Whereas expression of genes involved in recognition, e.g. *PGRP-SA* (GenBank Acc. No.: EFN70060, previously annotated as *PGRP-2*) and *GNBP* (EFN66519), and antimicrobial activity (e.g. *hymenoptaecin*, GenBank Acc. No. for EST: HS410972) is significantly lower in the midgut tissue especially in the pupal stages, two negative immune regulators, the amidase PGRPs *PGRP-LB* (EFN73971) and *PGRP-SC2* (EFN73970), are highly expressed in the midgut tissue in late pupae. The cleavage of *Blochmannia* PGN via the amidase activity of these PGRPs might reduce the amount of immunostimulating particles and therefore prevent an immune activity against the endosymbionts. Consequently, it was proposed that a down-modulation of the immune response might enable endosymbiont tolerance within the midgut tissue (Ratzka et al., 2013a).

Despite the broad and detailed knowledge about the dynamics of the symbiont distribution in the midgut and the symbionts' involvement in metabolism supporting their hosts, only little is known about the bacterial distribution during ant oogenesis. Former studies focussed on the localisation of structures, initially believed to be fungi, especially during embryogenesis of *Camponotus ligniperda* (Blochmann, 1882; Buchner, 1918; Hecht, 1923). Within the embryos these structures, which are now known to be bacteria of the genus *Blochmannia*, are located into developing bacteriocytes of the determined midgut tissue but also into cells which are closely positioned to the predestined ovarian tissue (Hecht, 1923). In the ovarian tissue itself, developing oocytes in fertile animals were observed

to be infected with *Blochmannia* via follicle cells (Blochmann, 1882; Buchner, 1918). Additionally, it was shown that matured oocytes are the only cells infected within ovaries of *C. floridanus* (Sauer et al., 2002; Schröder et al., 1996). Nevertheless, knowledge about the mode of action of vertical endosymbiont transmission in *C. floridanus* remained limited.

1.6   The aim of this work

As described in the previous chapter, earlier work suggested the involvement of the *C. floridanus* immune system in the tolerance and regulation of its bacterial endosymbiont *B. floridanus*. Current knowledge point towards a down-modulation of the innate immune response in the endosymbiont bearing midgut tissue, especially during the late pupal stage (Ratzka et al., 2013a). Several high throughput approaches such as transcriptome analyses were found to be suitable to shed light on immune components and their regulation within other insects (Bang et al., 2015; Colgan et al., 2011; Gandhe et al., 2006; Kolliopoulou et al., 2015; Li et al., 2015; Meng et al., 2015; Sackton et al., 2013; Vogel et al., 2011). In order to further characterise the *C. floridanus* immune system, a transcriptome analysis was performed (Gupta et al., 2015; Ratzka, 2012). In this work, the transcriptome data set of *C. floridanus* was further validated via quantitative RT-PCR and analysed in order to describe the general composition of the ants' innate immune system. Interestingly, there is also evidence that the insect immune response may differ between several developmental stages of the insects (Colgan et al., 2011; Fellous and Lazzaro, 2011; Randolt et al., 2008; Wilson-Rich et al., 2008). Accordingly, this work focussed on the question whether there are differences in immune-relevant gene expression patterns between different developmental stages of *C. floridanus*.

The insect haemolymph serves as transport medium for molecules and metabolites, immune cells and effectors, and plays a role in monitoring physiological conditions. In several studies the haemolymph proteomes of mostly *Drosophila* but also other insects have been analysed especially with regard to the involvement of proteins in development, metabolism, and immune response (de Morais Guedes et al., 2005; Feng et al., 2014; Karlsson et al., 2004; Levy et al., 2004; Vierstraete et al., 2004a; Vierstraete et al., 2004b; Woltedji et al., 2013; Zhang et al., 2014b). Here the proteomes of haemolymph collected from immune challenged and control *C. floridanus* adults and larvae was analysed focussing on the identification of immune-relevant proteins. Humoral effectors such as AMPs are produced mainly by the fat body and released into the haemolymph mostly in the later phase of an immune response. Thus, the expression kinetics of haemolymph peptides after immune challenge was analysed. The *C. floridanus* hymenoptaecin multipeptide precursor protein contains seven peptides with probable bioactivity. Differences in the amino acid sequences may result in structural and functional differences of the seven predicted peptides. Therefore, two *C. floridanus* hymenoptaecin peptides were produced via different recombinant expression systems and purified for functional antimicrobial assays.

The vertical transmission of *B. floridanus* is a key characteristic of the symbiosis and previous studies suggest that the endosymbionts are of highest importance during their hosts' ontogeny. Thus, the maternal transmission via oocytes or eggs enables the transfer of the endosymbionts to the next generation and therefore their hosts' complete development. Hence, the distribution of *B. floridanus* in ovaries of *C. floridanus* was intensively analysed with regard to cell types infected with the

endosymbionts during oogenesis. Also, possible transport mechanisms of *B. floridanus* were discussed. Since a down-modulation of the *C. floridanus* immune system was suggested to play a role in symbiont tolerance within the midgut tissue (Ratzka et al., 2013a), it seems very likely that immune-relevant genes might be regulated in the ants' ovaries as well. Accordingly and for the first time, the expression of immune-relevant genes within the ovarian tissue of *C. floridanus* was analysed.

2 Material

2.1 *Camponotus floridanus* colonies

For experiments performed in this work, different *C. floridanus* colonies, which are listed below (Table 1), were used.

Table 1 *C. floridanus* colonies used in this work. A.E. = Annett Endler, C-P.S. = Christoph-Peter Strehl

Colony	Place of origin	Laboratory rearing since	Collected by
C79	Florida, Orchid Island	31-08-2001	A.E., C-P.S.
C90	Florida, Orchid Island	31-08-2001	A.E., C-P.S.
C96	Florida, Orchid Island	31-08-2001	A.E., C-P.S.
C132	Florida, Sugarloaf Shores	05-07-2002	A.E.
C152	Florida, Orchid Island	09-07-2002	A.E.
C264	Florida, Tarpon Springs	23-06-2003	A.E.

2.2 Bacterial strains and cell lines

The bacterial strains used in this work are listed in Table 2.

Table 2 List of bacterial strains used in this work and their characteristics.

Bacterial strain	Genotype/ Characteristics	Reference
Escherichia coli DH5α	F$^-$, *endA1*, *glnV44*, *thi-1*, *recA1*, *relA1*, *gyrA96*, *deoR*, *nupG*, Φ80*dlacZ*ΔM15, Δ(lacZYA-argF)U169, *hsdR17(r_{K-} m_{K+})*, λ–	Bethesda Research Laboratories (BRL)
E. coli DH5α _pFBDM4_mutGST[TEV]*hld*(MCS1) _*egfp*(MCS2) 4	*E. coli* DH5α strain transformed with pFBDM4_mutGST[TEV]*hld*(MCS1) _*egfp*(MCS2)	this work
E. coli DH5α _pFBDM4_6xHis[TEV]*hrep*(MCS1) _*egfp*(MCS2)	*E. coli* DH5α strain transformed with pFBDM4_6xHis [TEV]*hld*(MCS1) _*egfp*(MCS2)	this work
E. coli DH10MultiBac	F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 recA1 endA1 araD139 Δ (ara, leu)7697 galU galK λ- rpsL nupG /pMON14272 Δ (chiA, vcath) / pMON7124	(Berger et al., 2004)
E. coli DH10MultiBac_GST[TEV]*hld*(MCS1) _*egfp*(MCS2) 76	*E. coli* DH10MultiBac strain transformed with pFBDM4_mutGST[TEV]*hld*(MCS1) _*egfp*(MCS2)	this work
E. coli DH10MultiBac_6xHis[TEV]*hrep*(MCS1) _*egfp*(MCS2) 2	*E. coli* DH10MultiBac strain transformed with pFBDM4_6xHis[TEV]*hrep*(MCS1) _*egfp*(MCS2)	this work
E. coli D31	F$^-$, *proA23*, *lac-28*, *tsx-81*, *trp-30*, *his-51*, *rpsL174*(strR), *rfe-229*, *ampCp-1,* Str^r, Amp^r	(Monner et al., 1971)
E. coli Rosetta 2(DE3)pRARE2	F$^-$, *ompT*, *hsdS_B(r_B^- m_B^-)*, *gal*, *dcm*(DE3)pRARE2(Cam^r); pRARE2	Novagen, Merck KGaA
E. coli Rosetta 2(DE3)pRARE2/pET15b_*hld*	*E. coli* Rosetta 2(DE3)pRARE strain transformed with pET15b_*hld;* Amp^r, Cam^r	(Kupper, 2011)

Bacterial strain	Genotype/ Characteristics	Reference
E. coli Rosetta 2(DE3)pRARE2/pET15b_*hrep*	*E. coli* Rosetta 2(DE3)pRARE strain transformed with pET15b_*hrep;* Amp^r, Cam^r	(Kupper, 2011)
E. coli Rosetta 2(DE3)pRARE2/pET32a_*hld* 3	*E. coli* Rosetta 2(DE3)pRARE strain transformed with pET32a_*hld;* Amp^r, Cam^r	this work
E. coli Rosetta 2(DE3)pRARE2/pET32a_*hrep* 2	*E. coli* Rosetta 2(DE3)pRARE strain transformed with pET32a_*hrep;* Amp^r, Cam^r	this work
Staphylococcus aureus	test strain	Kindly provided by Prof Vilcinskas (University of Gießen, Germany)

Additionally, eukaryotic cells *Sf*21 (*Spodoptera frugiperda*) (Smith et al., 1985; Vaughn et al., 1977) were used in adherent and suspension culture for recombinant expression of hymenoptaecin peptides. The insect cell line, which is a derivative of the *Sf*9 cell line, was kindly provided by Nils Neuenkirchen (Department of Biochemistry, Würzburg).

2.3 Oligonucleotides

Oligonucleotides were synthesised and purified either via desalting or HPLC by Sigma Life Science (Sigma-Aldrich Chemie GmbH, Munich, Germany)

2.3.1 Oligonucleotides for recombinant expression of hymenoptaecin derived peptides in different expression systems

Hymenoptaecin peptides were expressed in different expression systems. The following list presents the oligonucleotides used for several cloning strategies.

Table 3 Oligonucleotides used for cloning hymenoptaecin genes into different expression vector systems. (*hld* = hymenoptaecin-like domain; *hrep* = hymenoptaecin repeat domain)

Primer name	Sequence	T_m (°C)
Oligonucleotides for recombinant expression in bacteria:		
Cfl_*hrepNco*I_F1	TACCATGGATGCCGAAAAGGGTTCGAC	70
Cfl_*hrepNco*I_F1	CGCCATGGCTCAAGGAACTTCACTAAGCT	72
Cfl_*hymBamH*I_R1	TAGGATCCTTAGAAGCGGTAGCCAGCACTTA	72
pET T7 forward	TAATACGACTCACTATAGG	51
pET T7 reverse	GCTAGTTATTGCTCAGCGG	57
Oligonucleotides for production and sequencing of pFBDM4 vector constructs:		
Cfl_*hrepEcoR*I_F1	CGGAATT**C**CAAGGAACTTCACTAAGCT	67
Cfl_*hldNde*I_F1	TACATATGGATGCCGAAAAGGGTTCGAC	69
Cfl_*hymXba*I_R1	CCTCTAGATTAGAAGCGGTAGCCAGC	70
Cfl_*hrepNde*I_F2	CGCATATGCAAGGAACTTCACTAAGCT	67
pFBDM4_EGFP(MCS2)_mut*hrep*R2	TGCATATGGAATTCGCCCTGAAAATAAAG	66
EGFP(*Nco*I)_F1	GATTCCATGGTGAGCAAGGGCG	66
EGFP(*Not*I)_R1	CCGCGGCCGCTTATCTAGATCC	68
M13 fw	GTAAAACGACGGCCAGT	52
M13 R1	CAGGAAACAGCTATGAC	50
pFBDM4_mut_SeqO_F1	CAATGTGCCTGGATGCGTT	57
pFBDM4_mut_SeqO_F2	CTCAATTTCAACCACGCCCG	60

Primer name	Sequence	T_m (°C)
pFBDM4_mut_SeqO_F3	TTTCGCCCTTTGACGTTGG	57
pFBDM4_mut_SeqO_F4	GTTCTTTCCTGCGTTATCCCC	61
pFBDM4_mut_SeqO_F5	TGATGCCCATACTTGAGCCA	58
pFBDM4_mut_SeqO_F6	GCTGAACTTGTGGCCGTTTA	58
pFBDM4_mut_SeqO_F7	TGAAGCCATACCAAACGACG	58
pFBDM4_mut_SeqO_F8	TTGTTTGCCGGATCAAGAGC	58
pFBDM4_mut_SeqO_F9	ATCCTGCCAACTCCATGTGA	58
pFBDM4_mut_SeqO_F10	TAGATCTCACTACGCGGCTG	60
pFBDM4_mut_SeqO_F11	TTGTTCGCCCAGGACTCTAG	60
pFBDM4_mut_SeqO_R1	TTTGTCACATGGAGTTGGCAG	60
pFBDM4_mut_SeqO_R2	CGCGGAGTTGTTCGGTAAAT	58
pFBDM4_mut_SeqO_R3	CCCGAAAAGTGCCACCTAAA	58
MultiBac_MCS1_fwd	GGATTATTCATACCGTCCCA	49.7
MultiBac_MCS1_rev	CAAATGTGGTATGGCTGATT	47.7
MBacMCS2_F	GGACCTTTAATTCAACCCAACAC	61
MBacMCS2_R	ATTGTCTCCTTCCGTGTTTCAG	60

2.3.2 Oligonucleotides for validation of Illumina sequencing data via qRT-PCR

Oligonucleotide pairs for validation of Illumina sequencing results by qRT-PCR were designed on the corresponding mRNA sequences with Primer3 v. 0.4.0 (Rozen and Skaletsky, 2000) to yield products of 120-140 bp with melting temperatures around 56°C (Table 4).

Table 4 Oligonucleotides for validation of Illumina sequencing results via qRT-PCR

Gene name	GenBank Acc. No.	Primer name	Sequence
Housekeeping genes:			
60S ribosomal protein L32 (rpL32)	EFN68969	Cfl_rpL32_rtF1	GCCAACAGGCTTCCGTAAAG
		Cfl_rpL32_rtR1	TGAGCACGTTCGACGATAGC
Recognition:			
Alpha-2-macroglobulin-like protein 1 (TEP2)	EFN69033	Cfl_TEP1_cdsF1	ATATGGAGGCGTCATTTTCG
		Cfl_TEP1_cdsR1	CCTTCGACACCATCAGGA
CD109 antigen (TEP1)	EFN73645	Cfl_TEP2_cdsF1	ACGGTGGTATTGGCAGAC
		Cfl_TEP2_cdsR1	CTTGTGCAGGAATCCAAATG
CD109 antigen (TEP3)	EFN68790	Cfl_TEP3_cdsF2	TTACCGAGTACGGCAAATCC
		Cfl_TEP3_cdsR2	AGAAACCTCCGCAAAGTCAA
Signalling and immune defence:			
Nuclear factor NF-kappa-B p110 subunit; relish	EFN61437	Cfl_Rel_rtF1	CACCTTTGCAATTAGCTGCTG
		Cfl_Rel_rtR1	ACCTCCTTCGACTGCGATATG
Receptor-interacting serine/threonine-protein kinase 1 (ImdK1)	EFN61166	Cfl_ImdK1_cdsF1	ATGACGACGTTACGATGCAA
		Cfl_ImdK1_cdsR1	TATACTCTTTCGACTCGCTC
Hymenoptaecin (hym)	HS410972	Cfl_HymRT_F5	CACAGTAGAAACGAAAACATTCC
		Cfl_HymRT_R5	ATGAAGTTTCCTGGGCACTCG
NADPH--cytochrome P450 reductase nitric oxide synthase 1 (NOS1)	EFN67037	Cfl_NOS1_cdsF1	CCTTCCGATCCACCTTCT
		Cfl_NOS1_cdsR1	CTGAAAGGCGCTAGACCA
Mitogen-activated protein kinase kinase kinase 7 (MAPKKK7)	EFN63041	Cfl_MAPKKK7_cdsF1	TGGAGTCGTCTGGAAGGG
		Cfl_MAPKKK7_cdsR1	GGATGGGCAACTCTGGAT
Cathepsin D	EFN61010	Cfl_CathD_cdsF1	CCAACATCGACAACGTGATT
		Cfl_CathD_cdsR1	AACAGGATTTGTGGACTTTTTG
Embryonic polarity protein dorsal	EFN68841	Cfl_Dorsal_cdsF1	GTTACAATGGTCGCGCTAT
		Cfl_Dorsal_cdsR1	ACAAACACCTTGCTTGCACA

2.3.3 Oligonucleotides for gene expression analysis in ovarian tissue via qRT-PCR

The expression of putative immune genes in different tissues of orphaned *C. floridanus* workers was analysed by qRT-PCR. Oligonucleotide pairs for qRT-PCR were designed on the corresponding mRNA sequences with Primer3 v. 0.4.0 (Rozen and Skaletsky, 2000) to yield products of 120-140 bp with melting temperatures around 56°C (Table 5).

Table 5 Oligonucleotides used for qRT-PCR.

Gene name	GenBank Acc. No.	Primer name	Sequence
Housekeeping genes:			
Elongation factor-1 alpha (EF1a)	EFN72500	Cfl_EF1_rtF2	ACCCTTGGCGTTAAGCAGTT
		Cfl_EF1_rtR2	CCGGGTTGTAGCCAATCTTT
60S ribosomal protein L32 (rpL32)	EFN68969	Cfl_rpL32_rtF1	GCCAACAGGCTTCCGTAAAG
		Cfl_rpL32_rtR1	TGAGCACGTTCGACGATAGC
rpL18	EFN68908	Cfl_rpL18_rtF1	AAACCTGGACGGGAGAACTG
		Cfl_rpL18_rtR1	TGCTCGAGCTCTTTCGGTTA
glyceraldehyde-3-P dehydrogenase (GAPDH)	EFN69158	Cfl_GAPDH_rtF1	TTCACGACCATCGAGAAAGC
		Cfl_GAPDH_rtR1	TTGAAGCTCGGATCGTAAGC
Recognition:			
Peptidoglycan-recognition Protein SA (PGRP-SA)	EFN70060	Cfl_PGRP-2_rtF2	TGAGGGATGTGGATGGACAC
		Cfl_PGRP-2_rtR2	TTAATTGATGCGCAGCGTTC
PGRP-LB	EFN73971	Cfl_PGRP-LB_rtF2	ACAGCATGAATTTCGTGAATGC
		Cfl_PGRP-LB_rtR2	GCCATGATGGATCACCACATA
PGRP-LC	EFN63542	Cfl_PGRP-LE_rtF1	GAATTCTGCGAGACGCAATC
		Cfl_PGRP-LE_rtR1	CTGCCAATGTATGCGAGACC
PGRP-SC2	EFN73970	Cfl_PGRP-SC2_rtF2	AACTGATCGCTTATGGAGTGG
		Cfl_PGRP-SC2_rtR2	CCAATGAGGCCATGATTGTAT
beta-1,3-glucan-binding protein (GNBP)	EFN66519	Cfl_GNBP_rtF1	ATTTCCGGATAACACGACCA
		Cfl_GNBP_rtR1	CGCCGTTCCTATGTTTCCTA
beta-1,3-glucan-binding protein (GNBP1_3)	EFN62569	Cflo_GNBP1_3_rtF1	GCCGATTCCTTTAGAGACT
		Cflo_GNBP1_3_rtR1	TAAGACGTCCCGAAATCAC
		Cflo_GNBP1_3_rtF2	CGCGCCGCTTATAATAATGT
		Cflo_GNBP1_3_rtR2	TCGCTATTGTTCCTGCTTCT
Alpha-2-macroglobulin-like protein 1 (TEP2)	EFN69033	Cfl_TEP1_cdsF1	ATATGGAGGCGTCATTTTCG
		Cfl_TEP1_cdsR1	CCTTCGACACCATCAGGA
Serine proteinase:			
PREDICTED: *serine protease persephone (psh)*	XP_011251552 EFN71579.1	Cfl_SerPro2_rtF2	GTCACAGAGCATGCTGATCG
		Cfl_SerPro2_rtR2	ACCTTGCTCGCGTGAATTATC
Signalling:			
Nuclear factor NF-kappa-B p110 subunit; relish (rel)	EFN61437	Cfl_Rel_rtF1	CACCTTTGCAATTAGCTGCTG
		Cfl_Rel_rtR1	ACCTCCTTCGACTGCGATATG
Toll-interacting protein tollip	EFN63773	Cfl_Tollip_rtF1	GGATCAACAAGCTGCTCTGG
		Cfl_Tollip_rtR1	GGGTCCATCCTTGTCATTCC

2.4 Plasmids

The following Table 6 presents plasmid vectors used in this work.

Table 6 Plasmids used and generated in this work.

Plasmid	Description	Reference/ Supplier
Bacterial expression vectors:		
pET15b_*hld_Nde*I_*BamH*I	expression vector with fragment encoding *hld* in multiple cloning side, Amp^r	(Kupper, 2011)
pET15b_*hrep_Nde*I_*BamH*I	expression vector with fragment encoding *hrep* in multiple cloning side, Amp^r	(Kupper, 2011)
pET-32a	Bacterial expressions vector, N-terminal Trx•Tag™, His•Tag® and S•Tag™ for detection and purification of fusion proteins (cleavable), T7 promotor, *lac* operator, Amp^r	Novagen, Merck Biosciences, Merck KGaA, Darmstadt, Germany
pET-32a _*hld* 3_*Nco*I_*BamH*I	pET32a containing the gene encoding the hymenoptaecin-like peptide (hld)	this work
pET-32a _*hrep* 2_ *Nco*I_*BamH*I	pET32a containing the gene encoding a hymenoptaecin repeat domain (hrep)	this work
Baculovirus system transfer vectors:		
pFBDM4_GST[TEV](MCS1)_ *egfp*(MCS2)	MultiBac system: modified bacterial transfer vector, derivative of pFBDM (Berger et al., 2004), multiple cloning site MCS1 contains GST•Tag - cleavable by TEV protease (*Tobacco Etch Virus nuclear-inclusion-a endopeptidase*), MCS2 contains gene coding for enhanced green fluorescence protein (*egfp*)	(Neuenkirchen, 2012)
pFBDM4_mutGST[TEV]*hld* (MCS1)_*egfp*(MCS2)	Derivative of pFBDM4_GST[TEV](MCS1)_*egfp*(MCS2), contains altered TEV cleavage site (mut) and gene coding for hld	this work
pFBDM4_ 6xHis[TEV](MCS1)_(MCS2)	MultiBac system: modified bacterial transfer vector, derivative of pFBDM (Berger et al., 2004), multiple cloning site MCS1 contains 6xHis•Tag - cleavable by TEV protease, MCS2 contains *egfp*	(Neuenkirchen, 2012)
pFBDM4_ 6xHis[TEV]*hrep*(MCS1)_*egfp* (MCS2)	Derivative of pFBDM4_6xHis[TEV](MCS1)_(MCS2), contains gene coding for hrep and gene coding for egfp	this work

2.5 Antibodies

Antibodies, used for the detection of proteins via Western blot, are listed below (Table 7).

Table 7 List of antibodies for Western blots.

Antibody	Host	Dilution	Supplier
Primary antibodies			
His-probe Antibody (H-3)	mouse	1:500	Santa Cruz
αGST antibody	rabbit	1:100,000	Invitrogen™
6xHis-Hrep-probe Antibody	rabbit	1:1,000	this work, in cooperation with immunoGlobe
Secondary antibodies			
ECL^TM anti-mouse IgG HRP-linked	goat	1:2,000	Santa Cruz
ECL^TM anti-rabbit IgG HRP-linked	goat	1:2,000	Santa Cruz

2.6 Culture media, solutions and buffers

All chemicals used for production of media, buffers, and used for other solutions were supplied by Applichem, Bio-Rad, Eppendorf, GE Healthcare, Greiner, Invitrogen, Merck, Roche, Roth, Serva and Sigma-Aldrich.

2.6.1 Bacterial culture media and supplementary solutions

LB medium (Luria Broth) (agar):

10 g tryptone/ peptone
5 g yeast extract
10 g NaCl
(15 g agar)

➢ add 1 l dH$_2$O
➢ autoclave and add antibiotics, if necessary, when temperature < 50°C.

Table 8 List of antibiotics and solutions supplementing culture media and stock concentrations.

Antibiotic	Concentration	Solvent
Ampicillin	100 mg/ml	ddH$_2$0
Chloramphenicol	34 mg/ml	100% ethanol
Kanamycin	50 mg/ml	ddH$_2$0
Streptomycin	100 mg/ml	ddH$_2$0
Tetracycline	10 mg/ml	ddH$_2$0
Gentamycin	50 mg/ml	100% ethanol
Other solutions		
Isopropyl β-D-1-thiogalactopyranoside (IPTG)	1 M	ddH$_2$0
5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal)	2%	dimethylformamide

2.6.2 Buffers for Agarose gel electrophoresis, SDS-PAGE and Western blot

1x TBE (Tris/EDTA/borate)

89 mM Tris-HCl, pH 8.3
89 mM boric acid
2 mM Na$_2$EDTA

Agarose gel

0.5%-2.0% (w/v) agarose
1x TBE
➢ add HDGreen™ DNA-Dye (INTAS Science Imaging Instruments GmbH, Göttingen, Germany; 4-6 μl/ 100ml gel solution)

SDS-PAGE stacking gel – 5%

5% (v/v) acrylamide solution (30%)
0.125 M Tris-HCl, pH 6.8
0.001% (w/v) SDS
0.0005% (w/v) APS
0.0025% (v/v) TEMED

SDS-PAGE resolving gel – (12-16.5%)

12-16.5% (v/v) acrylamide solution (30%)
0.375 M Tris-HCl, pH 8.8
0.001% (w/v) SDS
0.0005% (w/v) APS
0.005% (v/v) TEMED

4x SDS-PAGE protein sample buffer (50ml)

10 ml Tris-HCl (2M), pH 6.8
15 ml SDS (10%)
20 ml glycerol
3.75 ml β-mercaptoethanol
50 mg bromophenol blue

SDS-PAGE buffer (10x)

30.25 g/l Tris
144 g/l glycine
10 g/l SDS

Coomassie staining solution

45% ethanol
10% acidic acid
➢ add Coomassie Brilliant Blue R250 (Serva Blue; 0.5 g/l)

Coomassie de-staining solution

10% (v/v) acetic acid

Western transfer buffer

5.8 g/l Tris
2.9 g/l glycine
0.37 g/l SDS
20% (v/v) methanol

10x PBS

1.37 M NaCl
27 mM Na_2HPO_4
20 mM KH_2PO_4

Western blot washing solution (1x PBS-T)

1x PBS
0.05% (v/v) Tween 20

ECL solution 1

100 mM Tris HCl pH 8.5
2.5 mM Luminol
0.4 mM p-coumaric acid

ECL solution 2

100 mM Tris HCl pH 8.5
0.02% H_2O_2

Stripping Buffer

7.8 g glycine
0.5 g SDS
1.54 g DTT
5 ml Tween
➢ ad 500 ml H_2O
➢ adjust pH to 2.2

DNA and protein markers used in this work: GeneRuler^{TM} 1kb DNA ladder (Fermentas, Life Technologies GmbH, Darmstadt, Germany), PageRuler™ Unstained Low Range Protein Ladder and PageRuler^{TM} Prestained Protein Ladder (Fermentas). DNA samples were loaded on agarose gels in 6x DNA sample loading buffer (Thermo Fisher Scientific).

2.6.3 Buffers for fluorescence *in situ* hybridisation (FISH)

4% Paraformaldehyde (PFA)

20 g PFA
480 ml PBS
➢ heat to 60°C and adjust pH to 7.2
➢ add 500 ml 1x PBS

FISH Hybridisation buffer

35% formamide
0.9 M NaCl
20 mM Tris-HCl pH 7.5
0.2% SDS

FISH wash buffer

70 mM NaCl
20 mM Tris-HCl pH 7.5
5 mM EDTA pH 8.0
0.01% SDS

2.6.4 Buffers for protein purification

Lysis Buffer immunoGlobe (iG)

50 mM Tris-HCl, pH 8.0
0.25% sucrose
1 mM EDTA

Detergent Buffer iG

20 mM Tris-HCl, pH 7.5
2 mM EDTA
200 mM NaCl
1% (w/v) deoxycholic acid
1% Nonidet P-40

Washing Buffer iG

0.5% Triton X-100
1 mM EDTA

GST-Resuspension/ Wash Buffer 01

50 mM Tris-HCl, pH 7.5
0.3 M NaCl

GST-Elution Buffer 01

Resuspension/ Wash Buffer
➢ 10 mM glutathione, pH 8.0

GST-Resuspension/ Wash Buffer 02

50 mM Tris-HCl, pH 7.5
0.15 M NaCl

GST-Elution Buffer 02

50 mM Tris
10-20 mM glutathione, pH 8.0

GST-Resuspension/ Wash Buffer 03

1xPBS, pH 7.4

GST-Elution Buffer 03

50 mM Tris-HCl
10-20 mM glutathione, pH 8.0

His-Resuspension/ Wash Buffer 01

20 mM Hepes-NaOH, pH 7.5
1 M NaCl
10% glycerol
5 mM β-mercaptoethanol
20 mM imidazole

His-Elution Buffer 01

20 mM Hepes-NaOH, pH 7.5
1 M NaCl
10% glycerol
5 mM β-mercaptoethanol
250 mM imidazole

His-Resuspension/ Wash Buffer 02

20 mM NaH_2PO_4
500 mM NaCl
20 mM imidazole

His-Elution Buffer 02

20 mM NaH_2PO_4
500 mM NaCl
750 mM imidazole

Binding Buffer (HisÄKTA)

20 mM NaH_2PO_4
500 mM NaCl
45 mM imidazole
pH 7.4

Elution Buffer (HisÄKTA)

20 mM NaH_2PO_4
500 mM NaCl
700 mM imidazole
pH 7.4

2.6.5   Buffers for various applications

Ant ringer solution (after M. Obermayer, (Kleineidam et al., 2005))

Solution A in 900 ml ddH_2O:
7.4 g NaCl
0.5 g KCl
0.22 g $CaCl_2$

Solution B in 100 ml ddH_2O:
0.11 g Na_2HPO_4
0.05 g KH_2PO_4

➢ mix solution A and B and then add:
- 1.1 g TES
- 1.2 g trehalose
➢ adjust to pH 7.0

3   Methods

3.1   Ant rearing

C. floridanus colonies were kept in artificial plaster nests in climate chambers at 25°C and a 12 hour light-dark cycle as described before (Feldhaar et al., 2007). Animals were fed twice or three times per week with cockroaches and honey water (50% w/w). For several experiments in the context of this work, the developmental stages of *C. floridanus* were defined as has been described before (Stoll et al., 2009). While very young larvae (L1) are less than 2 mm long and still clustered together, older larvae are about 2-4 mm long (L2). The pupal stages comprise P1 (pupae before metamorphosis), P2 (older pupae, still uncoloured, shortly after metamorphosis) and P3 (late pupae, slightly melanised, shortly before eclosion). Young workers (W1) are not fully melanised and show no aggressive behaviour yet. In contrast adult workers (W2) are completely melanised but show no additional features to further distinguish their exact age. Additionally, *C. floridanus* adult workers can be classified into two castes due to a distinct size polymorphism. In the following the two worker castes are referred to as minor and major workers.

For the characterisation of the localisation of *B. floridanus* within worker ovaries, orphaned/ queenless sub-colonies were generated. Like in ants in general, *C. floridanus* colonies also show a strict reproductive division of labour between queens and workers. In queenright *C. floridanus* colonies, intermediate levels of reproductive constraints have been observed (Khila and Abouheif, 2008). Worker ovaries, which are not fully sterile but reduced in size, become activated after workers are separated from the queen. As a consequence, workers produce unfertilised haploid eggs, which develop into males (Wolschin et al., 2004). In order to obtain egg laying workers, queenless sub-colonies consisting of 60 to 80 major workers and at least 200 minor workers were generated. To maintain nursing behaviour, brood was provided to the sub-colonies. In order to avoid suppression of worker reproduction by surface hydrocarbons on queen eggs (Endler et al., 2004), only older larvae (at least L2) and pupae P1-P3 were provided. After two to four months, eggs, larvae, and consequently adult males were produced by the orphaned workers indicating activation of worker ovaries.

Antimicrobial treatment of *C. floridanus* facilitates the removal of the endosymbionts from the midgut tissue (Zientz et al., 2006). To test the ovarian tissue for antibiotic accessibility and to characterise possible changes in *Blochmannia* infestation in ovaries, one orphaned sub-colony of *C. floridanus* (C90) with egg laying workers was divided into two parts of equal size. Larvae and eggs in both sub-colonies were removed and one sub-colony was fed normally with honey water (50% w/w) and cockroaches. The other colony was fed with cockroaches and honey water (50% w/w) containing 1-2% rifampicin (Carl Roth). The amount of eggs, larvae, pupae and adult males produced in both sub-colonies was analysed over the following eight weeks. Additionally, the distribution of *Blochmannia*

endosymbionts in the ovarian and midgut tissue was analysed via fluorescence *in situ* hybridisation (chapter 3.11.1).

3.2 Immune challenge of different *C. floridanus* stages

3.2.1 Production of bacterial suspensions

Depending on the experimental approach, different developmental stages of *C. floridanus* were immune challenged with heat-killed Gram-negative *E. coli* D31 and the Gram-positive *M. luteus*. To obtain stocks with defined bacterial concentrations, a single colony from the particular bacterial strain was transferred from a fresh agar plate into 6 ml LB medium and incubated at 37°C (30°C in case of *M. luteus*) and 190 rpm overnight. In case of *E. coli* D31 LB medium was supplemented with 100 µg/ml ampicillin and 100 µg/ml streptomycin. An aliquot of the bacterial overnight culture was taken for determination of the cell concentration and the residual culture was immediately autoclaved. A serial dilution of the aliquot was made and 100 µl of the dilutions 10^{-6} to 10^{-8} were each plated twice on LB agar plates and incubated at 37°C or 30°C overnight. Colonies per plate were counted, and the initial concentration of the bacterial cultures was calculated. The dead autoclaved bacteria from the initial culture were pelleted through centrifugation for 3 minutes (min) at 6,000 x g and washed twice with sterile 1xPBS. Bacterial pellets were directly used for pricking of *C. floridanus* or diluted with ant Ringer solution to the required concentration for injection experiments.

3.2.2 Immune challenge of *C. floridanus*

Different developmental stages of *C. floridanus* were immune-challenged either via injection or via pricking. Injection of major and minor workers (W1-W2) was performed as described before (Geier, 2009). In short, the solution for injection (0.3-0.5 µl) was gathered with a micropipette and applied from the micropipette tip into a pulled glass capillary, which was produced in a DMZ-Universal Puller (Zeitz-Instruments Vertrieb GmbH, Martinsried, Germany). The glass capillary was put in a plastic hollow cylinder placed in a micromanipulator (Science Products GmbH) for injection. Adult workers were anesthetised on ice and then fixed in a micropipette tip glued on a small petri dish. Using the micromanipulator, workers were carefully injected between the first and the second segment of the gaster. It is of importance that no liquid leaks from the puncture wound after injection. Larvae were injected carefully at the ventral side on the level of the midgut.

For immune challenge via pricking, a minutiae needle (Minutiennadel Sphinx V2A 0,1x12mm, bioform, Nuremberg, Germany) was glued into the opening of a cannula as a fixture for better handling. The needle was directly dipped into a bacterial pellet containing *M. luteus* and *E. coli* D31 in a 1:1 ratio and adult worker ants were pricked between the first and second segment of the gaster. As for injection, larvae were pricked ventrally on the level of the midgut.

3.3   RNA isolation

For total RNA extraction from insect tissue (midgut, ovary) anesthetised animals were dissected in PBS and the corresponding tissue (Figure 4) was immediately put into RNAlater® (Ambion), a reagent inactivating RNases and stabilising RNA within tissues. Tissues from the same treatment groups were pooled in 500-1000 µl RNAlater and stored at 4°C until RNA isolation (for maximal 3 days).
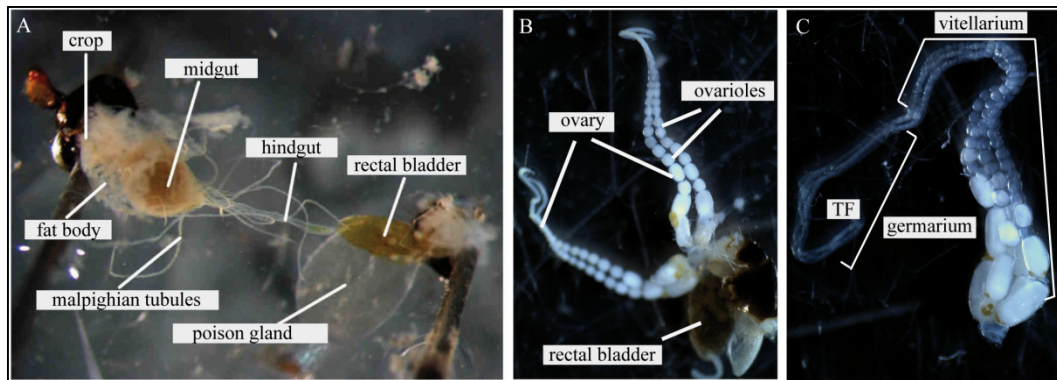
Figure 4 Internal organs of *C. floridanus*. **(A)** Anatomy of the intestinal tract of a *C. floridanus* worker (modified from (Stoll, 2009)). **(B)** A *C. floridanus* worker ovary consisting of two branches with two ovarioles each is shown. **(C)** The three main parts of each ovariole are the terminal filament (TF), the germarium, which harbours the germ-line stem cells and their progeny, and the vitellarium which comprises clusters of nurse cells and oocytes, termed egg chambers, in a linear arrangement. Within the vitellarium, oocytes undergo growth and vitellogenesis and finally enter maturation.

For total RNA extraction from whole larvae and whole workers or single organs and residual body parts, anesthetised individuals were dissected in RNase-free 1xPBS. In order to allow complete permeation of the tissue with RNAlater, the cuticles were carefully opened at the head, thorax, and abdomen. Additionally, the poison gland was removed and emptied into PBS before putting the animals into RNAlater. Up to 50 mg of tissue samples stabilised in RNAlater were transferred into 600 µl buffer RLT (RNeasy mini kit, Qiagen) supplemented with 6 µl β-mercaptoethanol within a 2 ml tube containing lysing matrix D (MP Biomedicals). Samples were homogenized within a FastPrep®-24 Instrument (MP Biomedicals) at 6.0 to 6.5 m/s for 45 seconds (sec) (4°C). Insoluble debris such as exoskeleton was removed by centrifugation at 14,000 x g for 3 min at 4°C. The supernatant was transferred into a new 2 ml microcentrifuge tube and 1 ml TRI Reagent® LS (Sigma-Aldrich) was added to each sample. The following steps of RNA isolation were performed as described in the manufacturer's protocols and each RNA pellet was finally resuspended in RNase-free water (100 µl). The RNA concentration and purity of each sample were measured with a NanoDrop™ ND-100 spectrophotometer. In case of high DNA contamination indicated by a 260/280 ratio below 2.0 as well as for samples experientially known to be highly DNA-contaminated (e.g. samples from whole larvae), RNA samples were further purified through RNeasy mini kit columns (Qiagen) with on-column DNase digestion (RNase-Free DNase Set, Qiagen), as described in the manufacturer's procedures. This additional purification promotes removal of contaminating protein or phenol residues, which was indicated through a 260/230 ratio below 2. Finally, the RNA was eluted in 20-40

µl RNase-free water. RNA concentration and purity of each sample were measured with a NanoDrop™ spectrophotometer. RNA samples further used in real-time quantitative reverse transcription polymerase chain reaction (qRT-PCR) analysis were additionally checked for gDNA contamination by PCR with specific oligonucleotides (about 200-500 ng RNA per PCR reaction). RNA samples were stored at -80°C until further use.

3.4   Haemolymph collection

For haemolymph sample collection, *C. floridanus* workers or larvae were anesthetised on ice. Larvae were pricked with a minutiae needle or a pulled glass capillary laterally directly behind the head and a micro capillary pipette (Hartenstein, Würzburg, Germany) was held closely to the wound while the animal was slightly pressed with forceps. In case of workers, head, thorax, and abdomen were separated using dissecting scissors. By applying slight pressure on head and thorax, the haemolymph was squeezed out of the body parts and collected with a micro capillary pipette. Haemolymph samples were held on ice and a few microliters of a mixture of aprotinin and N-phenylthiourea (each 0.2 mg/ml; Sigma-Aldrich) were added to avoid haemolymph melanisation. Samples were centrifuged at 14,000 x g for 1 min at 4°C to remove haemocytes and cellular debris. Finally, supernatants were collected and haemolymph samples were stored at -20°C (-80°C for long time storage).

3.5   Analysis of the immune defence inventory of *C. floridanus* by transcriptome sequencing

3.5.1   RNA sample preparation for Illumina sequencing

To gain insights into the immune system of *C. floridanus* and to unravel possible adaptations to the ants' lifestyle as social insects and hosts of the intracellular endosymbiont *B. floridanus*, the published genomic sequence (Bonasio et al., 2010) of the ant was re-evaluated and extended with genome-wide transcriptome data. The preparation of RNA samples for Illumina sequencing was performed by Carolin Ratzka (Ratzka, 2012). Briefly, a 1:1 mixture of autoclaved *E. coli* D31 and *M. luteus* (6-8*10^9 cells/ml) was prepared as described in chapter 3.2.1) and held on ice during the experiment. Adult minor workers (W2) and larvae (L2) were pricked with a minutiae needle dipped into the bacterial pellet. Afterwards, animals were kept in artificial nests for 12 hours (h). Non-inoculated animals were treated in parallel as controls. Total RNA was extracted from each five immune-challenged and five untreated L2 and W2 as described in chapter 3.3. RNA concentration and quality were determined on an Agilent 2100 Bioanalyzer using the Agilent RNA 6000 Nano Chip kit (Agilent Technologies, Böblingen, Germany) according to the manufacturer´s instructions.

Each 7.5 µg total RNA from immune-challenged workers and larvae as well as each 5 µg from untreated workers and larvae were mixed. The two resulting RNA samples (immune-challenged and control) were sent to Eurofins MWG GmbH (Ebersberg, Germany) for sequencing. An amplified short insert cDNA library of 150-250 bp size for each sample was prepared and each library was sequenced in 2 channels of Illumina HiSeq 2000 using 2x 50 bp paired-end sequencing (Gupta et al., 2015; Ratzka, 2012).

3.5.2 Bioinformatical analyses of the Illumina sequencing data set

The sequencing data obtained from Eurofins MWG GmbH was analysed by Frank Förster and Shishir Gupta (Department of Bioinformatics, University of Würzburg). Initially, the data set was applied to the Tuxedo pipeline including the TopHat and Cufflinks tools (Trapnell et al., 2012). These tools were used for the assembly of the expressed transcripts. The merged transcriptome annotation was generated using cuffmerge, and the differential gene expression analysis was performed using cuffdiff (v2.0.2) and, independently, by use of DESeq (Anders and Huber, 2010). The R package cummeRbund (v0.1.3, http://comp-bio.mit.edu/cummeRbund/) was used for the exploration of differentially expressed genes and for the visualisation of differently spliced genes. Furthermore, the sequencing data set was used to re-annotate the previously published genome sequence of *C. floridanus* (Bonasio et al., 2010) utilising an extended version of the Generalized Hidden Markov Model (GHMM) based *ab initio* predictor Augustus v2.7 (Stanke et al., 2006a; Stanke et al., 2006b; Stanke and Waack, 2003).

New gene annotations were used for immune pathway reconstructions. In short, interactions among the components of the Toll, Jak-Stat, IMD, the and JNK signalling pathway of *D. melanogaster* were mined individually from three pathway databases KEGG (Kanehisa et al., 2008), FlyReactome (http://fly.reactome.org/), and INOH (Yamamoto et al., 2011) followed by manual curations in reference to scientific literature. Proteins involved in the networks were chosen as queries for BlastP search (Altschul et al., 1997) against *C. floridanus* proteins. Additionally, a domain analysis with PFAM search (Finn et al., 2010) was performed and the identified homologous immune proteins of *C. floridanus* were mapped onto the reconstructed immune signalling networks of *D. melanogaster* (Gupta et al., 2015). Finally, the HMM search module of HMMER3 (Fjell et al., 2007) and BlastP search were used for identification of AMPs and other putative immune effectors.

3.5.3 Validation of Illumina sequencing data by qRT-PCR

Differential gene expression results obtained by transcriptome analysis were validated performing qRT-PCR. Therefore, *C. floridanus* minor workers (W2) and late larvae (L2) of six different ant colonies (C90, C96, C152, C79, C264, C132) were injected with heat-killed *M. luteus* and *E. coli* D31 via pricking. Twelve hours post injection RNA samples from each five W2 and five L2 were isolated as described above (chapter 3.3). For validation of Illumina sequencing data, RNA was not only

isolated from treated but also from untreated animals. Here, samples from workers and larvae were handled separately enabling analysis of differences in gene expression between the two developmental stages.

3.5.3.1   cDNA synthesis for qRT-PCR

For the validation of Illumina sequencing data and general gene expression studies 0.5-2.0 μg of total RNA were reverse-transcribed with oligo (dT)18 primers using the RevertAid First Strand cDNA Synthesis Kit (Fermentas) according to the manufacturer´s protocol. Resulting cDNA samples were purified with the QIAquick PCR purification kit (Qiagen, Hilden, Germany) and eluted in buffer EB (10 mM Tris HCl, pH 8.5) to a final concentration of 10 ng/μl. Produced cDNA samples were stored at -20°C until further use.

3.5.3.2   qRT-PCR using the ddCt method

Oligonucleotide pairs for qRT-PCR were designed on the corresponding mRNA sequences as mentioned above (chapter 2.3.2, Table 4). For gene expression studies, qRT-PCR experiments were performed using the PerfeCta™ SYBR® Green FastMix®, Rox™ (Quanta Biosciences, Gaithersburg, MD, USA) on a StepOnePlus™ Real-Time PCR System (Applied Biosystems). The PCR master mix contains the fluorescent dye SYBR Green I, which incorporates into the DNA produced during PCR. The fluorescence of each reaction batch is proportional to the generated amount of DNA and is measured after each cycle.

Samples for qRT-PCR contained 1x PerfeCta™ SYBR® Green FastMix®, Rox™, gene-specific oligonucleotides in a concentration of 250 nM each, 1 μl of the cDNA (10 μg/μl), and dH$_2$O to a final volume of 20 μl. The enzymes within a sample were initially activated for 5 min at 95°C. Subsequently, 45 cycles of 5 sec denaturation at 95°C, 10 sec of annealing at 56°C and 20 sec of extension at 60°C were run. Fragment specificity was checked in a melting curve. Each biological sample was run in duplicate and results were averaged. The relative transcription levels were calculated using the ddCt method.

The so-called Ct- (cycle threshold) value of each sample was determined for comparison of different samples. Ct-values represent the number of cycles at which the fluorescence significantly exceeds the background fluorescence and can be used for calculating the relative transcription level of one sample (treatment) in comparison to another sample (control). The method used here, is the so-called ddCt method (Livak and Schmittgen, 2001). In short, the Ct-value of a constitutively expressed reference gene (housekeeping gene) is subtracted from the Ct of the target gene. The resulting dCt-value represents internal standardisation between different samples. In a next step, the dCt-values from the control samples are subtracted from the treatment samples to gain ddCt-values. The ratio of the relative expression between control and treatment sample can be calculated by the equation 2^{-ddCt}. If no difference in gene expression occurs, the ratio is equal to one.

3.5.3.3 Reference gene selection

The stable expression of the reference gene under the chosen conditions is an absolutely necessary requirement for internal standardisation of target gene expression levels. In the course of this work, BestKeeper, an excel-based tool which determines the best suited housekeeping gene from up to ten candidate genes and combines them into an index, was used (Pfaffl et al., 2004).

For validation of Illumina sequencing data the suitability of the four reference genes *ribosomal protein L32* (*rpL32*), *ribosomal protein L18* (*rpL18*), *elongation factor 1-alpha* (*EF1α*) and *glyceraldehyde-3-phosphate dehydrogenase* (*GAPDH*) was determined using the BestKeeper software tool on the basis of their expression levels within different tissues (midgut and residual body parts) and developmental stages (L2, P3, W2) of *C. floridanus* (Ratzka, 2012). All four candidate genes were considered as suitable genes for normalization in *C. floridanus*. Thus, *rpL32* was used as an internal standard for this experiment.

3.5.3.4 Statistical analysis of qRT-PCR data regarding immune challenge in *C. floridanus*

Statistical analysis for gene expression studies in immune-challenged *C. floridanus* was performed using Statistica v10 enterprise x64. A two-sided T-test for matched pairs was used to test, whether the relative gene expression of six biological replicates of immune challenged animals (dCT immune) differs significantly from relative gene expression of control animals (dCT control). Gene expression was considered as being downregulated with a ratio < 0.5 and upregulated with a ratio > 2. Additionally, normalised changes in gene expression in the ants were tested for the influence of the two factors "treatment" (immune challenge (i)/ control (c)) and "developmental stage" (larvae (L)/ workers (W)) using factorial ANOVA followed by a Tukey´s HSD post hoc test.

3.6 Analysis of the *C. floridanus* haemolymph proteome

3.6.1 Preparation of haemolymph protein samples for MALDI-TOF

Haemolymph samples for Matrix-Assisted Laser Desorption-Ionisation time of flight (MALDI-TOF) mass spectrometry were prepared as follows. After immune challenge of *C. floridanus* larvae and adult workers (chapter 3.2), haemolymph samples were collected at three different time points (6 hours post injection (hpi), 24 hpi and 48 hpi) and haemocytes were removed via centrifugation with 300 x g for 10 min at 4 °C. For separation of peptides from proteins within the samples equal volumes of haemolymph from each time point were taken, mixed and acetone was applied to each sample in a 1:5 ratio. Mixed samples were incubated at -80 °C for 2 h. Final separation was achieved by centrifugation with 15,000 x g for 15 min at room temperature (RT). The supernatants containing peptides (< 8 kDa) were removed and the remaining protein samples (pellets) were provided to the cooperation partners

for MALDI-TOF mass spectrometry (Jens Vanselow, AG Schlosser, Rudolf-Virchow Centre, Würzburg).

Acetone precipitated proteins were solubilized in 1xLDS sample buffer (Thermofisher) with 50 mM DTT, denatured for 10 min at 95°C and cleared by centrifugation (10 min at 16,200 x g). Samples were alkylated with 120 mM iodoacetic acid and loaded onto a 4-12% Bis/Tris gel (Thermofisher). The gel was stained with SimplyBlue Safe Stain (Thermofisher) and each lane was cut into 15 pieces. The excised gel bands were de-stained with 30% acetonitrile in 100 mM ABC buffer, shrunk with 100% acetonitrile, and dried in a vacuum concentrator. Digests with 0.1 μg trypsin per gel band were performed in 100 mM ammonium bicarbonate buffer (ABC) overnight at 37°C. The digestion supernatants were combined with peptides extracted from the gel slices with 5% formic acid.

3.6.2   Mass spectrometry analysis

NanoLC-MS/MS analyses were performed on an LTQ-Orbitrap Velos Pro (Thermo Scientific) equipped with an EASY-Spray Ion Source and coupled to an EASY-nLC 1000 (Thermo Scientific). Peptides obtained by digestion of protein samples with trypsin were loaded on a trapping column (2 cm x 75 μm ID, PepMap C18, 3 μm particles, 100 Å pore size) and separated on an EASY-Spray column (25 cm x 75 μm ID, PepMap C18, 2 μm particles, 100 Å pore) with a 30 minute linear gradient from 3-30% acetonitrile and 0.1% formic acid with 200 nl/min. MS scans were acquired in the Orbitrap analyser with a resolution of 30,000 at m/z 400. MS/MS scans were acquired in the Orbitrap analyser with a resolution of 7,500 at m/z 400 using HCD fragmentation with 30% normalised collision energy. A TOP5 data-dependent MS/MS method was used; dynamic exclusion was applied with a repeat count of 1 and exclusion duration of 30 seconds; singly charged precursors were excluded from selection. Minimum signal threshold for precursor selection was set to 50,000. Predictive AGC was used to target a value of $1e^6$ for MS scans and $5e^4$ for MS/MS scans. Lock mass option was applied for internal calibration in all runs using background ions from protonated decamethylcyclopentasiloxane (m/z 371.10124).

3.6.3   Bioinformatical data analysis

Raw MS data files were processed and analysed with MaxQuant version 1.5.2.8 (Cox and Mann, 2008). The database search was performed with tryptic cleavage specificity and three allowed miscleavages against a combined database containing the UniProt reference proteome of *C. floridanus* (UniProt proteome ID: UP000000311, taxon ID: 104421 download date: 2015-03-10), the new genome annotation Cflo_New (available at http://camponotus.bioapps.biozentrum.uni-wuerzburg.de/), and additionally a database containing common MS contaminants. Protein identification was under control of the false-discovery rate (< 1% FDR on protein and peptide level) at least 1 razor/unique peptide required for identification. Variable modifications were acetylation (protein N-terminal), glutamine to pyro-Glu formation (N-terminal Q), amidation (any C-terminal), and oxidation (M).

Carbamidomethyl (C) was set as fixed modification. For protein quantitation of immune-challenged versus control animals, the LFQ intensities were used (Cox et al., 2014). Proteins with less than two identified razor/unique peptides were dismissed. Missing LFQ intensities in the control samples were imputed with values close to the baseline if intensities in the corresponding immune-challenged samples were present. Data imputation was performed with values from a standard normal distribution with a mean of the 5% quantile of the combined log10 transformed LFQ intensities and a standard deviation of 0.1. To identify proteins significantly over- or underrepresented in either the control or the immune-challenged samples, intensity-binned boxplots (300 proteins per bin) were calculated. Outliers of the 1.5-fold or the 3-fold interquartile range (IQR) were marked as significantly altered.

A second set of bioinformatical methods was applied by Shishir Gupta (Department for Bioinformatics, University of Würzburg). In a first step the haemolymph MS data were pre-processed and the identifiers of the proteins were made consistent, by mapping the *C. floridanus* re-annotation IDs over the old *C. floridanus* accession numbers. Furthermore, Gene Ontology (GO) annotation was performed with Blast2Go suite (https://www.blast2go.com/) for annotating the biological process, molecular function and cellular compartment of the proteins present in haemolymph data.

3.7   Analysis of *C. floridanus* haemolymph peptides

3.7.1   Preparation of haemolymph peptide samples for MALDI-TOF

Haemolymph samples for MALDI-TOF mass spectrometry were prepared as follows. After immune challenge of *C. floridanus* larvae and adult workers (chapter 3.2) haemolymph samples were collected at five different time points (6 hpi, 18hpi, 24 hpi and 48 hpi, 72 hpi) and haemocytes were removed via centrifugation with 300 x g for 10 min at 4°C. For separation of peptides from proteins within the samples, equal volumes of haemolymph from each time point were taken, mixed and acetone was applied to each sample in a 1:5 ratio. Mixed samples were incubated at -80 °C for 2 hours. Final separation was achieved by centrifugation with 15,000 x g for 15 min at RT. The supernatants containing peptides (< 8 kDa) were collected and provided to the cooperation partner for MALDI-TOF analysis (Jens Vanselow, AG Schlosser, Rudolf-Virchow Centre, Würzburg). The supernatants (peptides) were lyophilized (ALPHA 1-2 LD plus, Martin Christ, Germany), peptides were re-dissolved in 100 mM ABC buffer, reduced by addition of DTT (final concentration 50 mM) and alkylated by adding iodoacetic acid (final concentration 120 mM) and 30 min incubation at RT. Samples were acidified with 1% tri-fluoric acid (TFA) and desalted using C18 Stage tips (Rappsilber et al., 2003). Each Stage Tip was prepared with 3 disks of C18 Empore SPE Disks (3 M) in a 200 µl pipet tip. Finally, peptides were eluted with 60% acetonitrile in 0.1% formic acid, dried in a vacuum concentrator (Concentrator 5301, Eppendorf, Germany), and stored at -20°C. Prior to nanoLC-MS/MS analysis, peptide samples were dissolved in 2% acetonitrile/ 0.1% TFA.

3.7.2   Mass spectrometry analysis

NanoLC-MS/MS analyses were performed on an LTQ-Orbitrap Velos Pro (Thermo Scientific) equipped with an EASY-Spray Ion Source and coupled to an EASY-nLC 1000 (Thermo Scientific). Peptide samples were loaded on a trapping column (2 cm x 75 µm ID. PepMap C18, 3 µm particles, 100 Å pore size) and separated on an EASY-Spray column (either 25 cm x 75 µm ID, PepMap C18, 2 µm particles, 100 Å pore size or for haemolymph peptides in the first replicate on a 25 cm × 200 µm ID, PepSwift Monolith column). The PepMap C18 columns were run with a 30 minute linear gradient from 3-30% acetonitrile and 0.1% formic acid with 400 nl/min, while the monolith column was run with a 60 min linear gradient and 2ul/min.

MS scans were acquired in the Orbitrap analyser with a resolution of 30,000 at m/z 400, MS/MS scans were acquired in the Orbitrap analyser with a resolution of 7,500 at m/z 400 using HCD fragmentation with 30% normalised collision energy. A TOP5 data-dependent MS/MS method was used; dynamic exclusion was applied with a repeat count of 1 and exclusion duration of 30 seconds; singly charged precursors were excluded from selection. Minimum signal threshold for precursor selection was set to 50,000. Predictive AGC was used to target a value of $1e^6$ for MS scans and $5e^4$ for MS/MS scans. Lock mass option was applied for internal calibration in all runs using background ions from protonated decamethylcyclopentasiloxane (m/z 371.10124).

3.7.3   Data analysis

Raw MS data files were processed and analysed with MaxQuant version 1.5.2.8 (Cox and Mann, 2008). The database search was performed against a database containing the UniProt reference proteome of *Camponotus floridanus* (UniProt proteome ID: UP000000311, taxon ID: 104421 download date: 2015-03-10), the new genome annotation Cflo_New (available at http://camponotus.bioapps.biozentrum.uni-wuerzburg.de/), and additionally a database containing common MS contaminants. For acetone precipitation supernatants, the search was performed without enzyme specificity. Protein identification was under control of the false-discovery rate (< 1% FDR on protein and peptide level) and required at least 1 razor/unique peptide for identification. Variable modifications were acetylation (protein N-terminal), Gln to pyro-Glu formation (N-terminal Q), amidation (any C-terminal), and oxidation (M). Samples alkylated with iodoacetic acid were searched with a fixed carbamidomethyl (C) modification.

For time series analysis and data clustering, peptide intensities for each modification specific peptide were summed up for each time point from the evidence results file. Only those peptides with intensity values for at least 4 time points were considered. Intensity values were log10 transformed, quantile normalized and the log2-fold change calculated against the t=0 time point. The median average deviation as a robust estimator of variance for each time series was calculated and peptides were considered as regulated (differentially expressed), when at least a single ratio of a time series was

outside 2 times the median of all median average deviations. Finally, the temporal profiles were clustered according to regulation utilising fuzzy c-means clustering (Futschik and Carlisle, 2005).

3.8 General molecular cloning methods

3.8.1 Polymerase chain reaction (PCR) and agarose gel electrophoresis

DNA target sequences were specifically amplified through polymerase chain reaction (PCR) using polymerases with proof-reading function for purposes of sequencing and protein production. For those applications, high sequence accuracy was necessary and provided by the used polymerases, such as ReproFast DNA Polymerase (Genaxxon bioscience GmbH, Ulm, Germany). For general amplifications, such as colony PCR, MolTaq polymerase (Molzym, Bremen, Germany) was used according to the manufacturer's protocol. Briefly, 2.5 µl 10xPCR buffer, 0.125 µl MolTaq polymerase, and 0.5 µl 10mM dNTPs were used with 5-200 ng template DNA for a 25 µl reaction. A PCR program generally started with an initial denaturation step at 95°C for 3 min followed by up to 35 cycles with defined temperature changes, which were programmed in a Thermocycler T3 (Biometra, Göttingen, Germany). Each cycle consisted of a denaturation step (30 sec at 95°C) followed by the primer annealing (20 sec at a primer specific temperature; the lowest melting temperature of the used primer pair) and a DNA synthesis step (1 min per kilobase (kb) at 72°C). Subsequently, a final extension step of 5-10 min at 72°C was performed.

The correct length of the PCR products was checked by TBE agarose gel electrophoresis. Subsequent to the PCR, 5 µl of each PCR reaction were combined with 1 µl 6x loading dye and electrophoresed on a 1.2% agarose gel in 0.5x TBE buffer alongside quantitative DNA markers. DNA fragments were separated according to their size at 160 V for 25-40 min. TBE gels contained HDGreen™ DNA-Dye (INTAS Science Imaging Instruments GmbH) and were directly photographed under UV light.

3.8.2 Preparation of chemically competent *E. coli* cells

Chemically competent *E. coli* DH5α and *E. coli* DH10 cells were produced for cloning of different vector plasmids. A single colony from a fresh plate was transferred into 3 ml LB medium containing sufficient amounts of antibiotic if necessary and incubated at 37°C and 190 rpm overnight. On the next day 2 x 50 ml LB medium were inoculated each with 500 µl of the overnight culture and incubated at 37°C and 190 rpm until an optical density OD_{600} of 0.5. Then, cultures were held on ice for 15 min. The cells were centrifuged for 10 min at 5,000 rpm and 4°C and pellets were carefully resuspended in 10 ml ice-cold 0.1 M $CaCl_2$. After incubation for 30 min on ice, cells were pelleted at 4,000 rpm and 4°C for 10 min. Each pellet was resuspended in 2.8 ml sterile 0.1 M $CaCl_2$ with 20% glycerine. Aliquots of 200-250 µl were stored at -80°C.

3.8.3 Purification of PCR products

PCR products were purified using the GeneJET PCR purification kit (Fermentas) as described in the manufacturer's procedures. PCR products were eluted in 20-50 µl of elution buffer. DNA concentration and purity were checked using a NanoDrop™ spectrophotometer and the DNA was stored at -20°C.

3.8.4 Isolation and purification of plasmid DNA

For small scale isolation of plasmid DNA, mini-plasmid-preparations were performed using the AxyPrep™ Plasmid Miniprep kit (Axygen) as described in the manufacturers´ protocols. Plasmid DNA was isolated and purified from 4 ml overnight culture (supplemented with the corresponding antibiotics) and finally eluted in 30 µl elution buffer. For large scale isolation of plasmid DNA maxi-plasmid-preparations were performed using the Plasmid Maxi Kit (Qiagen, Hilden, Germany). Purification was done following the manufacturers' instructions and purified plasmids were eluted in 100 µl dH$_2$O. Plasmid DNA concentration and purity were measured photometrically and samples were stored at -20°C.

3.8.5 Cloning of DNA fragments into plasmid vectors and transformation into competent *E. coli* cells

DNA fragments obtained via PCR as well as the chosen plasmid vectors were digested with the respective restriction enzymes (Fermentas) in a convenient restriction buffer (Fermentas) for 2-4 h at 37°C according to the manufacturer's instructions. Afterwards, DNA products were purified and further digested with another enzyme if necessary. The digested plasmid vectors were dephosphorylated via FastAP (a thermosensitive alkaline phosphatase) in FastAP buffer (Fermentas). Digestion with restriction enzymes resulted in sticky-ended DNA fragments which were ligated overnight at 16°C using T4 DNA Ligase (Fermentas). Ligation was performed, as described in the manufacturer´s procedures with a 3-8 molar ratio of PCR product: vector. After overnight incubation the ligation reaction was mixed with 100 µl chemically competent *E. coli* cells and incubated for 20 min on ice. Cells were heat-shocked for 50 sec at 42°C and then immediately cooled on ice for 2 min. After addition of 900 µl LB medium the culture was incubated for 1.5 h at 37°C and 190 rpm. Afterwards, 50-100 µl of the transformation culture were spread on pre-warmed LB agar plates (supplemented with antibiotics if necessary). The plates were incubated at 37°C overnight.

3.8.6 DNA sequencing of purified PCR products and plasmid DNA

Purified PCR products or plasmid DNA were sequenced by Seqlab (Göttingen, Germany). About 250 ng per kb or about 1.2 µg of plasmid DNA were mixed with 20-30 pmol of a specific oligonucleotide. The samples with total reaction volumes of 15 µl were sent to Seqlab and the obtained sequences were analysed using BioEdit version 7.2.0, FastPCR (6.1.56 beta 3) and NCBI Blast tools.

3.9 Production of an antiserum for detection of *C. floridanus* hymenoptaecin peptides

3.9.1 Isolation and purification of inclusion bodies for immunisation

In a previous work, the gene encoding a hymenoptaecin repeat peptide (*hrep*) was amplified from genomic DNA of *C. floridanus*, cloned into a pET15b vector plasmid and transformed into an *E. coli* Rosetta 2(DE3) strain for overexpression of the polyhistidine-(6xHis) tagged peptide. It was determined that the expressed recombinant fusion proteins were located in the insoluble fraction, the so called inclusion bodies (Kupper, 2011). Therefore, the hymenoptaecin peptides (hrep) were purified following the protocol for inclusion body purification for immunisation by the company immunoGlobe Antikörpertechnik GmbH (Himmelstadt, Germany).

A single colony of *E. coli* Rosetta 2(DE3)pRARE2/pET15b_*hrep*10 was transferred into 500 ml LB medium supplemented with ampicillin (100 µg/ml) and chloramphenicol (34 µg/ml). The culture was incubated overnight at 37°C and 190 rpm. On the following day, the overnight culture was diluted 1:10 in LB medium (supplemented as before) and incubated under constant shaking at 190 rpm and 37°C to an OD_{600} of 0.5. Protein expression was induced by applying 0.1 µM IPTG and the bacterial culture was incubated for 5 h at 37°C and 190 rpm. Bacteria were pelleted by centrifugation at 3,000 x g and 4°C for 15 min and pellets were stored at -80°C after complete removal of medium residues.

The bacterial cell pellet was washed in ice-cold 1xPBS and resuspended in 80 ml cold Lysis Buffer iG (chapter 2.6.4) per 1.5 litre of bacterial culture. After vigorous mixing, another 20 ml Lysis Buffer iG and 200 mg lysozyme were added. The suspension was incubated on ice for 30 min and cells were disrupted by sonication on ice (10 min, 50% intensity; 50% duty cycle). The lysate was mixed with 200 ml Detergent Buffer iG and centrifuged for 10 min at 5,000 x g. After careful removal of the supernatant the pellet was washed in 250 ml Washing Buffer iG. Again, the suspension was centrifuged and the supernatant was carefully removed. This procedure was repeated until a tight pellet was obtained. Subsequently, the pellet was thoroughly washed in 250 ml 70% ethanol. After a final centrifugation step, the pellet was resuspended in a small volume (up to 5 ml) of freshly prepared ice-cold 1xPBS and sonicated for 3 min (50% intensity; 50% duty cycle). Protein solutions were stored at -20°C until further use.

3.9.2 Sodium dodecyl sulphate polyacrylamide gel electrophoresis for qualitative analysis of protein solutions

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) allows separation of proteins on the basis of their molecular weight. SDS-PAGEs were used for qualitative and quantitative analyses of protein samples. In a first step, proteins are heat denatured and negatively charged in a SDS containing buffer, so that they migrate towards the anode within an electric field. Gels with a high acrylamide concentration allow better separation of small proteins due to smaller gel pores. A

low acrylamide concentration is sufficient for separation of bigger proteins. The compositions of typical stacking and resolving gels are given in chapter 2.6.2. In this work, protein samples were generally mixed with 2x SDS-PAGE protein sample buffer in an 1:2 ratio or with 4x SDS-PAGE protein sample buffer in a ratio of 1:4. Protein samples were denatured for 5 min at 95°C and centrifuged (1 min 14,000 x g) to reduce agglutination due to cellular debris. The samples and a protein molecular size standard (PageRuler™ Prestained Protein Ladder; PageRuler™ Unstained Low Range Protein Ladder) were then loaded onto the stacking gels and electrophoresed in 1x SDS-PAGE buffer at 100-160 Volts. After electrophoresis proteins were visualised via staining for 30-60 min in Coomassie staining solution and de-staining of the background in Coomassie de-staining solution overnight. Finally, protein gels were scanned with an ImageScanner III (GE Healthcare, Munich, Germany).

For better separation of especially small proteins, samples were applied to 15% acrylamide gels at 25 mA using Tricine-SDS-PAGEs (Schägger, 2006; Schägger and von Jagow, 1987). For this application, the cathode buffer was composed of 0.1 M Tris, 0.1 M Tricine (pH 8.3) and 0.1% SDS, whereas the anode buffer consisted of 0.2 M Tris-HCl (pH 8.9) (Schägger, 2006).

3.9.3   Measuring protein concentrations

The concentration of a protein solution was determined either by performing a Bradford assay or by using the measuring tool of ImageJ 1.49k. The latter method enables measuring the concentration of a single protein band on a protein gel, whereas the Bradford assay is sufficient for measuring total protein concentration of a solution.

Both methods require a calibration curve. For the Bradford assay a serial dilution of bovine serum albumin (BSA) was set up in a total volume of 800 µl each (1 µg/ml, 2 µg/ml, 4 µg/ml, 8 µg/ml and 10 µg/ml). Every sample was prepared in duplicate. The target protein samples were diluted 1000-fold in dH$_2$O in a total volume of 800 µl. Each sample was mixed with 200 µl of Bradford reagent Roti-Quant (Carl Roth, Karlsruhe, Germany) and kept at RT in darkness for 15 min. The Bradford reagent contains a triphenylmethane dye (Coomassie Brilliant Blue G-250) which binds to proteins within a sample causing a shift in its maximum absorption from 465 nm to 595 nm. The absorption maxima were measured with an Ultrospec 2100pro and 3100pro spectrophotometer (Amersham) and protein concentrations were calculated.

For the protein gel based method, a SDS-PAGE was prepared and the respective target samples as well as samples for a calibration curve were separated on the gel. The gel was scanned with ImageScanner III (GE Healthcare, Munich, Germany) and further analysed with ImageJ. The measuring tool allows determining the concentration of a single protein band. The picture was inverted and a rectangular box was drawn around a protein band. Pressing "Ctrl" and "M" simultaneously opens a table showing numbers representing light points counted within the rectangular. Dragging the

rectangular on a position close by, without a protein band, allows measuring light points of the background and therefore, background subtraction from the measurement for each protein band. After matching the measurements to defined amounts of protein within the calibration curve target protein band concentrations are calculated.

3.9.4 Immunisation with hymenoptaecin peptides

The concentration of the purified hymenoptaecin peptide solution (hrep solution) was measured with the Bradford assay and six aliquots of each 250 µl hrep solution (1 mg/ml) were sent to immunoGlobe, where a lab reared rabbit was immunised with the hrep antigen after taking a pre-immune serum. The antigens were applied in three boosts every two weeks. After six weeks, a small test serum was taken and the quality of the antiserum was analysed via Western Blot. In case of sufficient quality of the serum, two large scale volume serum samples were taken from the rabbit for laboratory use.

3.9.5 Qualitative analysis of produced antibody sera via Western Blot

The abundance of proteins within a given sample as well as the quality of an antiserum can be analysed with Western blot.

After separation of proteins within a sample via SDS-PAGE, proteins were transferred to a polyvinylidene fluoride (PVDF) membrane via semi-dry Western blotting. The membrane and 14 layers of Whatman paper (Blotting Paper 703, 0.38 mm thickness) were cut to the required gel size. The PVDF-membrane was activated through immersion in 100% methanol for 1 min. Afterwards, the membrane was washed in 1x Western transfer buffer. Meanwhile, the gel was removed from the glass plates, washed in dH$_2$O and equilibrated in 1x Western transfer buffer. The Whatman papers were equilibrated in 1x Western transfer buffer and the assembly of the Western blot was performed as follows (starting from the bottom (anode) to the top (cathode)): eight layers of Whatman paper, the PVDF-membrane, the gel and again six layers of Whatman paper. Air bubbles in between layers were carefully removed and the blotting chamber was gently closed. The transfer was performed at 100-150 mA for 2 hours.

After protein transfer, membranes were blocked using 5% milk/PBS-T (1xPBS-T supplemented with 5% milk powder) for 1 h. Membranes were briefly rinsed in 1xPBS-T and subsequently incubated over night with the indicated primary antibodies (e.g. the hrep-antiserum) at 4°C under constant but slow shaking. The following day, membranes were washed thrice for 10 min each in 1xPBS-T, incubated with the respective secondary antibody conjugated to horseradish peroxidase (HRP) for 1 h at RT, followed again by three washing steps in 1xPBS-T. Detection of the protein of interest was finalised by using chemiluminescent ECL solutions and carried out with a ChemoStar Imager v.0.2.37 (Intas® Science Imaging Instruments GmbH, Göttingen, Germany). For detection of several but different proteins within the same sample, membranes were incubated in stripping buffer (at 65°C for

30 min) to remove primary and secondary antibodies. Analysis of Western blots was performed using ImageJ.

3.9.6 Identification of proteins via mass spectrometry

In order to identify single proteins within a complex sample, mass spectrometry can be applied following the separation of proteins via SDS-PAGE. In this work, either haemolymph samples or samples from protein purifications were applied to appropriate acryl amid gels and proteins were separated. The Coomassie stained gels were carefully destained and bands chosen for mass spectrometry analysis were cut out with a fresh scalpel and transferred to a fresh eppendorf tube with ddH_2O. The samples were sent for protein identification via mass spectrometry, which was performed by Knut Büttner and colleagues (University Greifswald).

3.10 Expression and purification of hymenoptaecin peptides for functional analyses

3.10.1 Expression of hymenoptaecin peptides in insect cells

Baculoviridae are a family of large enveloped DNA viruses that are infectious only to arthropods, especially insects within the order Lepidoptera. The *Autographa californica* Multicapsid Nuclear Polyhedrosis Virus (AcMNPV) is one of the most commonly used Baculoviridae for studies in insect cultures and exhibits two virion phenotypes at different times and locations in the viral infection cycle (Figure 5). Via fusion at the plasma membrane of host midgut epithelium cells, the so-called occlusion derived virions (ODV) are taken up and the viral DNA is transported inside the nucleus, where it replicates. The so-called budded virions (BV) are produced in the late phase of the infection cycle via budding from the surface of infected cells. The envelope of BVs is derived from the plasma membrane and they are produced for secondary infection of deeper tissues. In the very late phase, viral nucleocapsids become enveloped within the nucleus to form new ODVs. Subsequently, ODVs are occluded within a crystalline occlusion matrix protein, the Polyhedrin. Finally, the infectious occlusion bodies containing the ODVs are released via cell lysis (Blissard, 1996; Murphy et al., 2004).
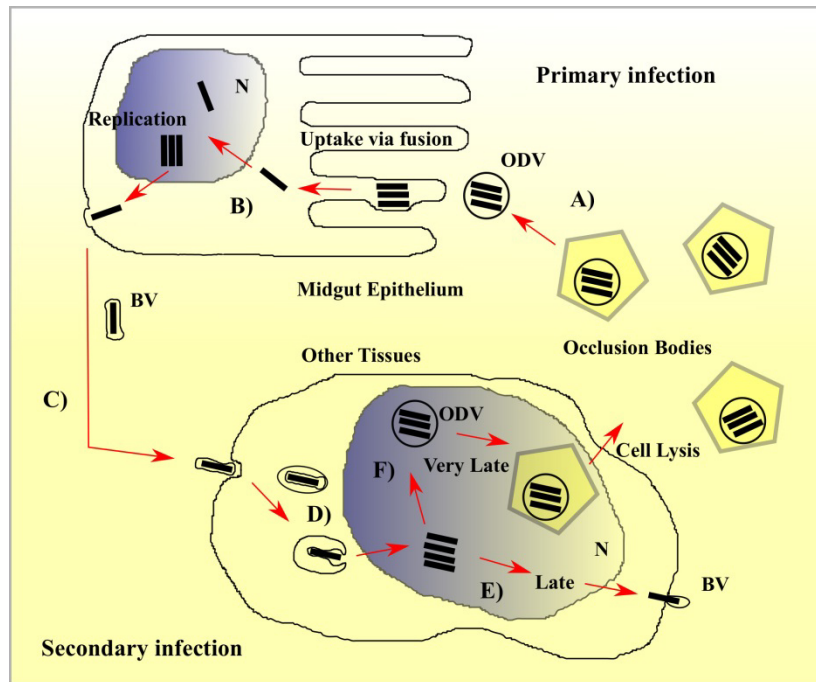
Figure 5 Baculovirus life cycle. A) Midgut epithelial cells of an insect host are infected by the occlusion derived virus (ODV) by fusion of the ODV at the plasma membrane. B) The viral genome is transported to the nucleus, where a round of replication occurs. C) Progeny virions subsequently bud from the basal side of the cell and budded virions (BV) enter other cells by endocytosis (secondary infection). D) Virions are released from endosomes and nucleocapsids are then transported to the nucleus, where gene expression, DNA replication, and assembly of progeny nucleocapsids occur. E) During the late phase, nucleocapsids are transported to the plasma membrane, where they bud to form BV again. F) During the very late phase of infection, nucleocapsids are enveloped within the nucleus and are occluded within a crystalline matrix of Polyhedrin. Occlusion bodies containing OVDs are released into the environment upon lysis of the cell.

For the baculovirus expression vector system (BEVS), the hyper-expression of the very late phase gene *polyhedrin* is exchanged for the chosen target gene. A gene of interest is cloned into a donor plasmid vector containing the *polyhedrin* promotor region. The recombinant donor plasmid is then transformed into competent *E. coli* DH10MultiBac cells (Berger et al., 2004). These cells harbour the Baculovirus genome (bacmid) as well as a helper plasmid providing the T7 transposase. Successful transposition is tested via antibiotic selection and blue/ white screening. The recombinant bacmid DNA is purified and then used for transfection of insect cells. The number of baculovirus particles was amplified by further infection of insect cells with the viral stock of passage P1. The baculovirus titre in each virus stock is determined by end-point dilution (O'Reilly et al., 1993). For large scale protein expression, insect cells are infected with a defined ratio of infectious agents to infection targets, the so-called multiplicity of infection (MOI) (Figure 6).
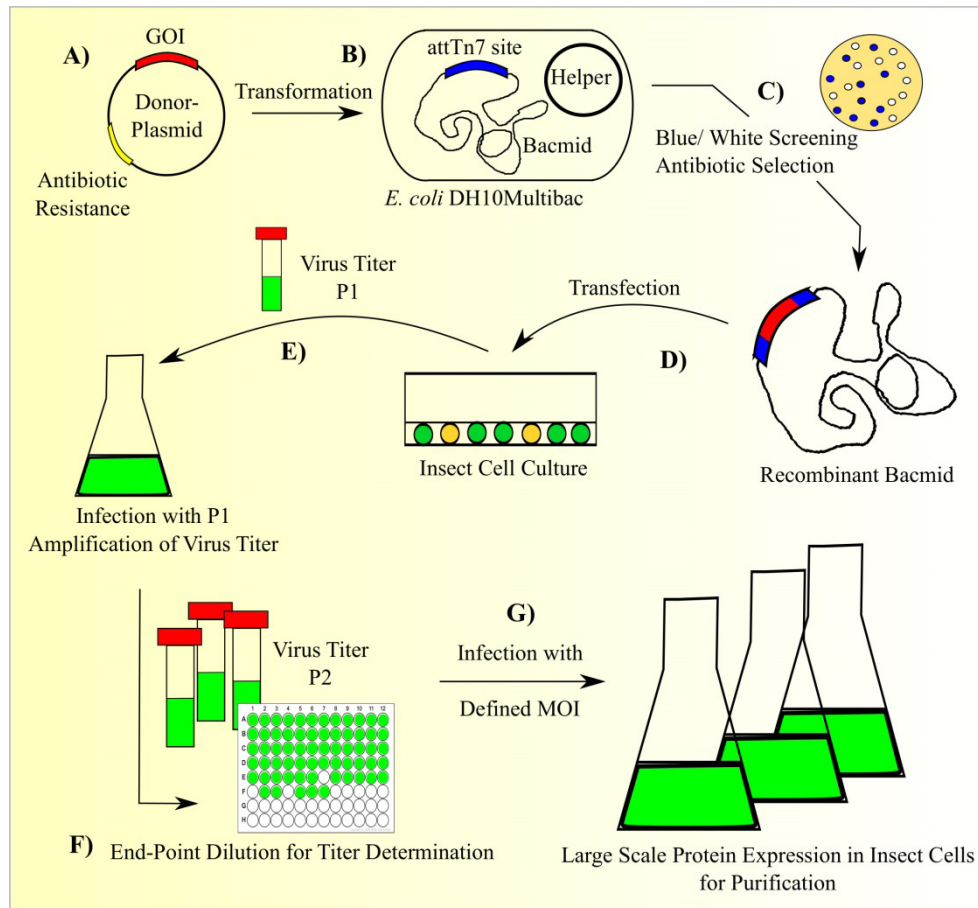
Figure 6 Scheme of the baculovirus expression vector system (BEVS) in insect cells. A) The genes of interest are cloned into a donor plasmid, which is then B) transformed into competent *E. coli* DH10MultiBac cells. C) Positive constructs are analysed by antibiotic selection and blue/ white screening. D) Recombinant bacmid DNA is then isolated and used to transfect insect cells to generate the initial virus stock P1. E) The number of baculovirus particles was amplified from passage 1 to passage 2 and F) the titer of each virus stock of passage P2 is determined by end-point dilution using the *egfp* reporter gene to mark positively transfected or infected cells (green). G) With the exact number of infectious particles defined, reproducible infections can be achieved for large scale protein expression.

The MultiBac system is a BEVS specifically designed for eukaryotic multiprotein expression (Berger et al., 2004; Fitzgerald et al., 2006). The pFBDM (plasmid FastBacDual-derived MultiBac) allows the incorporation of the gene expression cassette containing the gene of interest (GOI) via the Tn7 transposase. The vector contains two multiple cloning sites (MCS1 and MCS2). Whilst the MCS1 is enclosed by the polyhedrin promoter and the SV40 terminator, the MCS2 is encompassed by the p10 promoter and the HSVtk terminator. Derivatives of the plasmid vector pFBDM4, which contain modified multiple cloning sites MCS1 and MCS2 more suitable for further cloning, were used for the expression of hymenoptaecin peptides (Neuenkirchen, 2012). In this work, two plasmids enabling two different affinity purifications were chosen for expression of hymenoptaecin peptides in insect cells. The plasmid pFBDM4_GST[TEV](MCS1)_*egfp*(MCS2) contains a glutathione-S-transferase- (GST-) tag incorporated into the MCS1 for affinity purification and a gene encoding the enhanced green fluorescent protein EGFP as a transfection marker. The transfer vector pFBDM4_6xHis[TEV](MCS1)_(MCS2) contains a 6xHis-tag for affinity purification but no *egfp*

gene (Neuenkirchen, 2012). Both affinity tags are linked to target proteins via the specific cleavage side for the Tobacco Etch Virus nuclear-inclusion-a endopeptidase (TEV).

3.10.1.1 Incorporation of *hymenoptaecin* genes in transfer vectors and modification of TEV cleavage sites

Two DNA fragments encoding the putative hymenoptaecin peptides hymenoptaecin-like peptide (*hld*) and hymenoptaecin repeat peptide (*hrep*) were amplified using the oligonucleotide pairs Cfl_*hld*NdeI_F1 (for GST fusion proteins) and Cfl_hym*Xba*I_R1 as well as Cfl_*hrepEcoR*I (for 6xHis fusion proteins) and Cfl_hym*Xba*I_R1. *C. floridanus* adult workers were injected with heat-killed Gram-negative and Gram-positive bacteria as described before (chapter 3.2) and RNA samples were isolated (chapter 3.3). The respective cDNA samples were prepared (chapter 3.5.3.1) and used as templates for PCR with ReproFast DNA Polymerase.

The produced *Nde*I-*Xba*I-DNA fragments as well as the plasmid vector pFBDM4_ GST[TEV](MCS1)_*egfp*(MCS2) were digested with *Nde*I (Fermentas) in restriction buffer EB (Fermentas) for 2-4 h at 37°C according to the manufacturer's instructions. Afterwards, DNA products were purified as described before (chapter 3.8.3) and further digested with *Xba*I in 1x Tango buffer (both Fermentas). After incubation for 2 h at 37°C, DNA samples were purified and digested plasmid vectors were dephosphorylated via FastAP in FastAP buffer (Fermentas). The plasmid vector pFBDM4_ 6xHis[TEV](MCS1)_(MCS2) and *EcoR*I-*Xba*I-DNA fragments were digested simultaneously with both *EcoR*I and *Xba*I in 1x Tango buffer (Fermentas) for 2 h at 37°C. Digestion with restriction enzymes resulted in sticky-ended DNA fragments which were ligated overnight at 16°C using T4 DNA Ligase (Fermentas).

Additionally, a modification was introduced into the TEV cleavage sites encoded by the plasmid pFBDM4_GST*hld*[TEV](MCS1)_EGFP(MCS2). The plasmid was re-amplified completely using the oligonucleotide pair Cfl_*hld*NdeI_F2 and pFBDM4_EGFP(MCS2)_mut*hld*R2. Both reverse oligonucleotides comprise the codon GGC encoding glycine instead of the original codon CCG encoding proline. Introducing the GGC codon via side specific mutagenesis would result in plasmids encoding the correct and specific TEV cleavage sides (ENLYFQG). All oligonucleotides contain restriction sites for *Nde*I and the resulting PCR products were purified and digested with *Nde*I as described above.

Finally, DNA fragments encoding the transfection marker EGFP were incorporated into the multiple cloning site MCS2 of pFBDM4_6xHis*hrep*[TEV](MCS1)_(MCS2). The DNA fragment coding for EGFP was amplified using the oligonucleotides EGFP(*Nco*I)_F1 and EGFP(*Not*I)_R1 and the vector plasmid pFBDM4_GST*hld*[TEV](MCS1)_EGFP(MCS2) as a matrix. After purification and digestion with the restriction enzymes *Nco*I and *Not*I the *egfp* fragments were ligated with the likewise digested

plasmids pFBDM4_6xHis*hrep*[TEV](MCS1) _(MCS2). All DNA sequences of produced plasmid constructs (Table 6) were verified using the oligonucleotides listed in Table 3.

3.10.1.2 Transformation of *E. coli* DH10MultiBac cells and blue/white screening

Chemically competent *E. coli* DH10MultiBac cells were prepared as described above. These *E. coli* cells contain the unmodified viral bacmid DNA with an ampicillin and a kanamycin resistance gene. Additionally the strain contains a helper plasmid encoding the T7 transposase for recombination and a tetracycline resistance gene (Berger et al., 2004).

Eighty microliters of chemically competent *E. coli* DH10MultiBac were thawed on ice, incubated with 15 µg of recombinant transfer vector for 20 min, transferred to 42°C for 50 sec, placed on ice for 2 min, supplemented with 500 µl LB medium and incubated at 37°C shaking at 190 rpm for overnight. Cells were transferred onto agar plates containing kanamycin (50 µg/ml), 7 µg/ml gentamycin, 100 µg/ml ampicillin, 10 µg/ml tetracycline, 179 µg/ml IPTG as well as 32 µg/ml X-Gal and were incubated for 24-48 h at 37°C. The successful recombination was indicated by white colonies in the blue/ white screening. In blue colonies, the β-galactosidase transcribed by the *lacZ* gene cleaves X-Gal to yield galactose and 5-bromo-4-chloro-3-hydroxyindole. Oxidation of the latter results in 5,5'-dibromo-4,4'-dichloro-indigo, an insoluble blue product. Successful integration of the recombinant DNA disrupts the *lacZ* gene on the bacmid DNA and results in white colonies.

3.10.1.3 Isolation of recombinant bacmid DNA and analysis via PCR

Successfully transformed bacterial colonies containing recombinant bacmid DNA showed white colour and were used to inoculate 25 ml of LB medium supplemented with 50 µg/ml kanamycin, 7 µg/ml gentamycin, 100 µg/ml ampicillin, and 10 µg/ml tetracycline. Cultures were incubated at 190 rpm and 37°C overnight. Following centrifugation at 4,500 x g for 10 min at 4°C, the supernatant was carefully removed and the bacterial cell pellets were resuspended in 2 ml resuspension buffer S1 (50 mM Tris/HCl, 10 mM EDTA, (pH 8.0); 100 µg/ml RNase A). Aliquots of each 500 µl were prepared and the bacterial cells were lysed by addition of 500 µl of buffer S2 (200 mM NaOH, 1% (w/v) SDS) to each aliquot. Samples were carefully mixed and incubated for 5 min at RT. Dropwise addition of 500 µl ice-cold buffer S3 (2.8 M potassium acetate, pH 5.1) and 10 min incubation on ice caused the neutralisation of the solution and the precipitation of potassium dodecyl sulfate in combination with cellular debris. The bacmid DNA-containing fraction was separated by centrifugation for 10 min at 20,000 x g and RT. Equal amounts of the supernatant and isopropanol were combined, incubated on ice for 10 min and centrifuged for 60 min at 13,000 x g and 4°C. Finally, the supernatant was removed carefully and the bacmid DNA was washed with 375 µl 70% (v/v) ethanol (p.a.), air-dried under a sterile hood and dissolved in 300 µl 10 mM Tris/HCl (pH 7.4) overnight. Thus bacmid DNA preparations could be used for PCR analysis as well as insect cell transfection.

In addition to the described blue/ white screening, the successful recombination and correct insertion of a specific target DNA fragment was verified by PCR screens which were performed as described before (chapter 3.8.1). The purified bacmid DNA served as a template and one oligonucleotide binding to the inserted DNA sequence (gene specific primer) as well as an Oligonucleotide specific for the bacmid DNA (M13 primer) were applied to the reaction. The PCR reaction should only give positive results when transformation and recombination of the gene of interest into the bacmid DNA was successful. Respective products were purified and analysed via sequencing.

3.10.1.4 Propagation and transfection of insect cells

Insect cells (*Sf21*: *Spodoptera frugiperda*) were seeded at 0.5 x 10^6 cells/ml in the appropriate EX-CELL® TiterHigh™ culture medium (Sigma-Aldrich) and incubated in suspension culture at 27°C and 100 rpm using shaker flasks without baffles. The cell density and viability was measured using a haemocytometers (Neubauer improved) and applying trypan blue staining. The insect cell culture was mixed with 0.4% (w/v) trypan blue in 1xPBS in a 1:10 ratio and immediately applied to a haemocytometer. Whereas viable cells have intact membranes and exclude the dye, dead cells readily take up the blue dye through disrupted membranes and thus appear blue under the microscope. This method can therefore be used to identify viable cells. Cell cultures with a viability of at least 95% were used for virus stock generation and recombinant protein expression.

Once the cell density exceeded 3 x 10^6 cells/ml cells were diluted to 0.5 x 10^6 cells/ml. After 20 propagations (approximately 3 months), the *Sf21* cells were replaced by fresh ones. In order to prepare a backup of uninfected insect cells, exponentially growing cells were centrifuged, resuspended in 10% (v/v) DMSO, 45% (v/v) culture supernatant and 45% (v/v) fresh culture medium, aliquoted as samples of 2.5 x 10^7 cells in 1 ml and transferred to cryotubes. Insect cell stocks were cooled down slowly in a -80°C freezer at first and transferred to the vaporous phase of a liquid nitrogen storage container, later. For starting a new culture from frozen insect cells a vial (2.5 x 10^7 cells) was thawed rapidly, cells were washed with 10 ml of culture medium to remove any DMSO residues and finally seeded into 50 ml of fresh EX-CELL® TiterHigh™ medium at a concentration of 0.5 x 10^6 cells/ml.

To define the sufficient ratio of bacmid DNA to transfection reagent for *Sf*21 transfection trial experiments were performed in 6-well plates. Various amounts of bacmid DNA (1-10 µg) were tested in a combination with different amounts of the transfection reagent polyethylenimin (PEI; 1 mg/ml stock solution in dH₂O; Polyscience; (Ehrhardt et al., 2006)). For 1 µg bacmid DNA (about 130 kb) 0.65 µg of PEI were applied for a ratio of 1:5. Initially, 0.5 x 10^6 cells were seeded per well and left in the dark at 27°C to settle down. DNA and PEI samples were prepared separately in 250 µl EX-CELL® TiterHigh™ medium each. Samples were incubated at RT for 30 min, mixed carefully and incubated again for 30 min. The culture medium was removed from the cells and 1.5 ml of fresh medium were applied. Each transfection mixture was added dropwise. The transfection mixture was exchanged by 3 ml of fresh medium after either 5 hours or after overnight incubation. The cells were

further incubated at 27°C for 9-10 days. Since all recombinant baculoviruses in this study carried a reporter gene for the enhanced green fluorescent protein (EGFP), successful transfection was monitored by fluorescence microscopy.

The baculovirus stock P1 was generated by transfection of *Sf*21 cells with a pre-defined sufficient ratio of bacmid DNA and PEI. At first, 3×10^6 cells were seeded into a $75cm^2$ tissue culture flask and kept at 27°C for settling down. Transfection mixtures were prepared in each 1.2 ml EX-CELL® TiterHigh™ medium and incubated at 27°C for 30 min. The reactions contained 45 µg bacmid DNA encoding GST[TEV]hld in a ratio of 1:15 with PEI, or 45 µg bacmid DNA encoding His[TEV]hrep in a ratio of 1:5 with PEI. The culture medium was exchanged by 3 ml of fresh medium and the transfection solution before the cells were incubated at 27°C for 5 h. Eventually, the transfection solution was replaced by 12 ml of fresh EX-CELL® TiterHigh™ medium und incubated for 9-10 days at 27°C. Successful transfection was monitored by fluorescence microscopy and the culture supernatant containing produced infectious virus particles was collected as virus stock passage P1.

3.10.1.5 Amplification of baculoviruses and determination of the number of infectious virus particles by end-point dilution

After the transfection of the *Sf*21 cells with recombinant bacmid DNA, the cells generate infectious viral particles which are delivered to the culture medium by budding. These baculoviruses can be used for further infection of insect cells in order to increase the amount of viruses. These amplification steps last 72-79 hours instead of the initial generation of virus stocks after transfection. For the generation of the P2 baculovirus stocks, either the entire supernatant or 6 ml or the complete P1 stock were transferred to 100 ml uninfected *Sf*21 cells in suspension culture (2.0×10^6 cells/ml; viability > 95%) and incubated for 72-79 h at 27°C. Afterwards, insect cells were spun down via centrifugation and supernatants were carefully taken. These contain viral particles and can be stored at 4°C for up to 6 months.

The number of infectious viral particles was determined by end-point dilution (O'Reilly et al., 1993). The experiment was performed as described by O'Reilly *et al.* with some modifications. It was taken advantage of the fact that all recombinant baculoviruses express the EGFP. Ten-fold serial dilutions (10^{-1} to 10^{-8}) of the baculovirus stock were prepared in each 200 µl and uninfected insect cells were diluted to a concentration of 0.5×10^6 cells/ml. Then, 900 µl of the cell suspension were added to each dilution of the baculovirus stock. 60 µl of the mixtures were seeded into a 96-well plate. In total, twelve replicates of each virus dilution were added in one row of the 96-well plate. To avoid evaporation of the medium the plates were wrapped in plastic foil and incubation at 27°C for 9-10 days. After the incubation, each well was checked for infection and wells containing at least one cell expressing EGFP were scored as positively infected. The number of infectious particles per volume was determined by the method of Reed and Muench (1983), assuming that infected cultures would have been infected at any higher and all uninfected cultures would not have been infected at any

lower virus concentration (Reed and Muench, 1938). From this, the actual virus titre was calculated using a pre-set Excel spreadsheet (Appendix I, Table 19).

3.10.1.6 Protein expression in insect cells

For reproducible protein expression, the number of baculovirus particles per insect cell (multiplicity of infection = MOI) served as a major criterion for protein expression. Uninfected and exponentially growing *Sf*21 insect cells were pelleted at 600 x g for 5 min at RT, resuspended in fresh EX-CELL® TiterHigh™ medium and diluted to a final concentration of 1 x 10^6 cells/ml. Expression cultures were prepared in 2,000 ml shaker flasks without baffles containing culture volumes of 200-500 ml. Recombinant baculovirus was added to obtain a multiplicity of infection (MOI) of 2. The insect cells were incubated depending on the type of recombinant protein to be expressed for 48-79 hours at 27°C and 100 rpm and were finally harvested by centrifugation at 4°C and 600 x g for 30 min. The expression of proteins was observed via SDS-PAGE and Western blot.

3.10.1.7 Purification of GST-tagged hymenoptaecin-like peptide fusion proteins

For purification of GST-tagged hld, insect cell pellets from 1.0 l of expression culture were resuspended in 100 ml Resuspension Buffer supplemented with 4 ml Proteinase inhibitor cocktail (Roche, Mannheim, Germany). For this work different buffer combinations were tested (chapter 2.6.4). The cells were lysed by dounce-homogenisation and sonication (5 cycles of 30 sec pulse and 30 sec cooling on ice using a Sonifier 250 (Branson; output: 6, duty-cycle: 50%)). The cell suspension was supplemented with Nonidet (1%) and incubated on ice for 10 min. Afterwards samples were centrifuged for 60 min at 10,000 x g and 4°C. For protein binding, clear supernatant were incubated with 4 ml of pre-washed and equilibrated glutathione sepharose 4B slurry (GE Healthcare) per 1 litre of expression culture. After overnight incubation at 4°C, the flow-through was collected and proteins bound to the sepharose beads were washed with 5 column volumes (CVs) of the respective resuspension/ wash buffer. Finally, proteins were eluted in elution buffer containing 10-20 mM glutathione.

The protein fractions were quantitatively and qualitatively analysed via SDS-PAGE and Bradford assay. The eluted fractions containing GST-hld fusion proteins were pooled and supplemented with TEV protease (1 Unit TEV/ 3 µg fusion protein) and dialysed against the Digestion Buffer (50 mM Tris-HCl, pH 8.0; 0.5 mM EDTA; 0.5 mM DTT, pH 8.0) for 2 days at 4°C. TEV-cleavage separates the N-terminal affinity tag as well as the protease recognition site leaving an extra glycine residue on the remaining protein sequence (Kapust et al., 2002; Phan et al., 2002).

3.10.1.8 Purification of 6xHis-tagged hymenoptaecin peptide fusion proteins

For purification of 6xHis-tagged hymenoptaecin peptide fusion proteins insect cell pellets were resuspended in 100 ml of the respective Resuspension Buffers (chapter 2.6.4) supplemented with 4 ml

Proteinase inhibitor cocktail (Roche, Mannheim, Germany) per litre of expression cell culture. Insect cells were disrupted as described before. The cell lysate was cleared by centrifugation at 10,000 x g and 4°C for 30 min followed by another 30 min of centrifugation at 20,000 x g and 4°C. The cell lysate was subsequently incubated with 4 ml of Ni-NTA agarose beads (Qiagen) per litre culture by head-over-tail rotation overnight. After the beads were washed with 5 CVs of washing buffer, bound proteins were eluted by the addition of the respective elution buffer containing different amounts of imidazole. The quality of the protein samples was analysed via SDS-PAGE.

3.10.1.9 Cytobuster™ protein extraction of soluble proteins in insect cells

In addition to the purification procedures which were described before, GST- and 6xHis-tagged fusion proteins, were also purified with the Cytobuster™ protein extraction reagent (Novagen, An Affiliate of Merck KGaA, Darmstadt, Germany). The reagent is a formulation of detergents optimised for efficient extraction of soluble proteins from insect cells without applying mechanical treatment for cell disruption. Briefly, *Sf*21 cell pellets were washed in 5 ml 1xPBS and then resuspended in 150 µl Cytobuster™ reagent per 10^6 cells. The suspension was incubated for 5 min at RT before clearing via centrifugation for 5 min at 16,000 x g and 4 °C. The supernatant should contain soluble proteins, whereas the pellet should contain insoluble proteins and cellular debris. The quality of protein fractions was analysed via SDS-PAGE and Western blot.

3.10.2 Expression of hymenoptaecin peptides in *E. coli*

3.10.2.1 Incorporation of *hymenoptaecin* genes in the pET32a plasmid vector

In order to express recombinant hymenoptaecin peptides in bacterial cell cultures, two DNA fragments encoding the hymenoptaecin-like peptide (*hld*) and the hymenoptaecin peptide derived from one hymenoptaecin repeat domain (*hrep*) were amplified using the oligonucleotide pairs Cfl_*hldNco*I_F1 and Cfl_*hymBamH*I_R2 as well as Cfl_*hrepNco*I and Cfl_*hymBamH*I_R2. The pET plasmids (pET15b_*hld_Nde*I_*BamH*I or pET15b_*hrep_Nde*I_*BamH*I; (Kupper, 2011)) containing the respective DNA fragments encoding hymenoptaecin peptides were used as templates for PCR with ReproFast DNA Polymerase. The resulting *Nco*I-*BamH*I-DNA fragments as well as the plasmid vector pET32a were digested simultaneously with both *Nco*I and *Xba*I in 1x Tango buffer (Fermentas) for 2 h at 37°C. The produced sticky-ended DNA fragments were ligated for 2 h at RT using T4 DNA Ligase (Fermentas).

3.10.2.2 Transformation of competent *E. coli* Rosetta 2(DE3) cells

Chemically competent *E. coli* Rosetta 2(DE3) cells were prepared as described before. Eighty microliters of chemically competent *E. coli* Rosetta 2(DE3) were thawed on ice and incubated with 10-20 µl of ligation reactions for 20 min. The samples were transferred to 42°C for 50 sec, placed on ice for 2 min, supplemented with 500 µl LB medium and incubated at 37°C shaking at 190 rpm for

2 hours. The bacterial cells were transferred onto agar plates containing 30 μg/ml kanamycin and 30 μg/ml chloramphenicol. The plates were incubated at 37°C for 24 h. Transformed bacteria were selected via incubation on antibiotic selection plates and further analysed via colony-PCR with gene specific oligonucleotides.

3.10.2.3 BugBuster® protein extraction of soluble proteins in *E. coli* Rosetta 2(DE3)

The solubility of overexpressed proteins in *E. coli* cultures was determined via BugBuster® protein extraction following the manufacturer's instructions. Briefly, protein expression in a culture of either *E. coli* Rosetta 2(DE3)pET32a _*hld* 3 or *E. coli* Rosetta 2(DE3)pET32a _*hrep* 2 of mid-logarithmic growth (OD_{600} = 0.5) was induced with 0.1 mM IPTG. After the cultures were incubated for 1-3 h at 37°C and 190 rpm cells were pelleted by centrifugation for 10 min at 4,500 x g. The cell pellets were resuspended in 5 ml BugBuster® reagent per 1.0 g pellet. The suspension was supplemented with 1 μl Benzonase (25 Units) per ml BugBuster® reagent and incubated applying slow shaking at RT. After 20 min, insoluble proteins were separated from soluble proteins via centrifugation (20 min, 4°C and 16,000 x g). Small samples were taken from the supernatant (soluble fraction) for further analysis via SDS-PAGE. The pellet was again resuspended in 5 ml BugBuster® reagent per 1.0 g original bacteria pellet and lysozyme (200 μg/ml) was added. The suspension was mixed thoroughly and incubated for 5 min at RT. Six volumes of BugBuster® reagent diluted 1:10 in dH_2O were mixed with the sample followed by centrifugation for 15 min at 4°C and 16,000 x g. The supernatant was carefully removed and the washing step was repeated trice. Finally, the pellet containing insoluble proteins was resuspended in 1% SDS and a small sample was taken for further analysis with SDS-PAGE.

3.10.2.4 Large-scale overexpression of recombinant proteins in *E. coli* Rosetta 2(DE3)

The genes encoding hymenoptaecin peptides were incorporated into the pET32 plasmids. Thus, the expression of the respective fusion proteins is regulated by an IPTG inducible lac-promotor. The hymenoptaecin peptides were fused to thioredoxin-S-His$_6$ tags to increase protein solubility and therefore adapted for affinity purification. For large scale protein overexpression in bacterial cell cultures an overnight culture of *E. coli* Rosetta 2(DE3) harbouring either pET32a_*hld* 3 or pET32a_*hrep* 2 was diluted 1:10 in up to 5 l fresh LB medium. The bacterial cultures were incubated at 37°C and 190 rpm to an OD_{600} of 0.5. Control samples of each culture were taken and cells were pelleted via centrifugation for 3 min at 6,000 x g. Subsequently, cell pellets were resuspended in 2x SDS-PAGE protein sample buffer and stored at -20°C for SDS-PAGE analysis. The expression of proteins in bacterial cultures was induced with 1 mM IPTG and cultures were further incubated at 37°C and 190 rpm for 3 hours. Again, samples were taken and prepared for SDS-PAGE analysis. The remaining culture was centrifuged for 10 minutes at 4,500 x g and 4°C. Bacterial pellets were stored at -80°C until protein purification.

3.10.2.5 6xHis-tag purification of hymenoptaecin fusion proteins with the ÄKTA pure chromatography system

Recombinant hymenoptaecin peptide fusion proteins expressed in *E. coli* Rosetta 2(DE3) strains were purified via Immobilised Metal ion Affinity Chromatography (IMAC). For the subsequent high-resolution purification, an ÄKTA pure™ system (GE Healthcare) and a column pre-packed with nickel-nitrilotriacetic acid (NTA) sepharose beads was used. The affinity purification method is based on the binding of histidine residues to nickel-chelates of the column. Bound proteins can be eluted by either decreasing pH or by increasing concentrations of imidazole. Briefly, a cell pellet was thawed on ice and resuspended in Binding Buffer HisÄKTA at 5-10 ml per gram wet weight. Lysozyme was added to a concentration of 1 mg/ml and incubated on ice for 30 min. Cells were disrupted by six times sonication on ice for 10 sec (duty cycle: 50%; output: 5). Cellular debris was pelleted by centrifugation for 30 min at 10,000 x g and 4 °C. The supernatant containing soluble proteins was filtrated through a 0.45 µm filter. Only a cleared lysate can be applied to the nickel-NTA column.

Before sample loading, the column was washed with dH$_2$O and equilibrated with Binding Buffer HisÄKTA. In order to reduce unspecific protein binding, the column was then washed intensely with 7.5 column volumes of Binding Buffer. For protein elution, the Elution Buffer HisÄKTA containing imidazole was applied in a linear gradient to a final concentration of 100%. Samples of all steps (loading, wash and elution) were collected in fractions and further analysed via SDS-PAGE.

3.10.2.6 Proteolytic removal of the fusion tag via Enterokinase

The pET32a vector system allows the expression of thioredoxin-S-His$_6$ (Trx-S-His$_6$) tagged hymenoptaecin fusion proteins. Thus, the affinity tags can be removed via proteolytic cleavage with Enterokinase (GenScript, Piscataway, USA). This serine protease cleaves the peptide sequence Val-(Asp)$_5$-Lys. The N-terminal fusion tag is removed by cleaving after the fifth aspartate residue leaving the lysine residue at the N-terminus of the protein of interest. Following a small-scale optimisation fusion proteins were digested with either 50 Units Enterokinase per mg Trx- S-His$_6$-hld or 20 Units Enterokinase per mg Trx- S-His$_6$-hrep. Fusion proteins were digested in 1x Cleavage Buffer (GenScript) and in a volume sufficient to decrease salt and imidazole concentrations which are generally high in eluted protein fractions from His$_6$ tag purification.

Finally, protein samples were dialysed against a buffer containing 20 mM NaH$_2$PO$_4$ (pH 6.8). Dialysis was performed over 2 days at 4°C and solutions were exchanged for fresh buffer twice.

3.10.2.7 Ion-exchange chromatography and concentration of hymenoptaecin peptide samples

The ion-exchange chromatography is an appropriate method for separation of target proteins from complex samples. In this work, a cation exchange was used for the further purification of hymenoptaecin peptide solutions. Both putative antimicrobial peptides exhibit a very high isoelectric point (pI), which is defined as the pH at which a molecule does not carry a net electrical charge. Due

to this characteristic, the peptides should bind to a negatively charged column matrix (Capto S column, GE Healthcare) when dissolved within a buffer with pH < pI of the respective peptide.

Here, protein samples were loaded within a buffer containing 20 mM NaH_2PO_4 (pH 6.8). To minimise unspecific binding, the column was washed with 5 CV of binding buffer. Afterwards, proteins were gradually eluted with elution buffer (20 mM NaH_2PO_4, 1 M NaCl, pH 6.8). Samples of each fraction were taken and analysed via SDS-PAGE. Finally, fractions containing the target peptides were dialysed and concentrated in Dialysis buffer (50 mM Tris, 0.15 M NaCl, pH 7.5) via Amicon Ultra filter tubes (Merck Millipore, Germany).

3.10.2.8 Antimicrobial Activity Assay

Hymenoptaecin peptide samples were applied to 384-well plate based antimicrobial activity assays. The basic principle is the recording of growth curves of bacteria inoculated in liquid cultures in the presence of hymenoptaecin peptides or in untreated controls. Therefore, 10 µl of LB medium were placed in each well chosen for sample analysis. The hymenoptaecin peptide samples were applied in serial dilutions by adding 10 µl of the peptide solution to the first well and repeatedly transferring 10 µl of the mixture into the next well. Control samples were generated by adding pure buffer in the same manner as the peptide solutions. Finally, 10 µl of bacterial culture diluted to a starting OD_{600} of 0.001; which ensures full contact of the bacteria and peptides, were added to each well containing a defined LB-peptide solution mixture. The samples were carefully mixed and air bubbles were removed via a short centrifugation step. The 384-well plates were put into a Tecan Infinite® 200 PRO multimode reader and OD_{600} was measured in intervals of 15 min by applying constant orbital shaking with an amplitude of 3mm. Data were analysed with the Tecan i-control software, 1.10.4.0 and Microsoft Excel.

3.11 Analysing the localisation of *B. floridanus* within ovarian tissue of *C. floridanus*

3.11.1 Fluorescence *in situ* hybridisation

To analyse the distribution of *B. floridanus* in worker ovaries, fluorescence *in situ* hybridisation (FISH) was performed according to previously described protocols (Feldhaar et al., 2007; Stoll, 2009). Endosymbionts were visualised with a probe Bfl172 targeting a specific region of the *B. floridanus* 16S rRNA (5'-CCTATCTGGGTTCATCCAATGGCATAAGGC-3'). The probes were labelled with the fluorophore Alexa488 (Sigma-Aldrich Chemie GmbH, Munich, Germany). The respective sense probe Bfl172sense (5'-GCCTTATGCCATTGGATGAACCCAGATAGG-3'; labelled with Alexa488) was used as a negative control for specificity.

Ovaries of workers were dissected in ice-cold 1xPBS (pH7.4) and only ovaries with at least one matured egg at the posterior end were transferred in small glass vials for fixation in 4% (w/v)

paraformaldehyde in 1xPBS. After 2 hours of fixation at RT the fixative was exchanged by 1xPBS and samples were washed for five minutes. The washing step was repeated twice. Afterwards, samples were dehydrated by incubation in 30%, 50%, 70%, 90% and two times in 100% ethanol, each time 3 to 5 minutes. Hybridisation was performed in hybridisation buffer (35% formamide, 900 mM NaCl, 20 mM Tris/HCl pH 7.5, 0.2% SDS) with 500 mM of each oligonucleotide probe. Samples were placed in a humid chamber and incubated for 2-3 h at 46°C in the dark. After hybridisation, the samples were washed for 30 min in washing buffer (70 mM NaCl, 20 mM Tris/HCl pH 7.5, 5 mM EDTA, 0.01% SDS) at 48°C. In addition, DAPI was supplied to the wash buffer for counter staining of eukaryotic and bacterial DNA. Subsequently, samples were rinsed in sterile distilled water, placed on microscopy slides, mounted in Mowiol and analysed by confocal laser scanning microscopy (CLSM) with a Leica TCS SP5 (Leica Microsystems, Wetzlar, Germany). Pictures were taken with LAS AF version 2.7.3.9723 (Leica) and further processed using ImageJ 1.49v and Adobe Illustrator CS4 14.0.0.

3.11.2 Sample preparation for transmission electron microscopy and tomography

Ovaries of *C. floridanus* workers were dissected in ice-cold 1xPBS, separated from surrounding fat body tissue and transferred into Karnovsky's fixative (2% paraformaldehyde, 2.5% glutaraldehyde in 0.1 M PBS, pH 7.4) and left at 4°C for overnight fixation. Samples were washed five times for 5 min in cacodylic acid buffer (50 mM cacodylic acid, pH 7.2) and again fixed on ice in 2% osmium tetroxide in cacodylic acid buffer for 2h. After five times washing in dH$_2$O, samples were stained in 0.5% uranyl acetate overnight at 4°C. Staining was followed by dehydration in the presence of increasing ethanol concentrations (30 min each in 50, 70, 90% and two times in 100% ethanol at 4°C; two times 30 min in 100% ethanol at RT) and incubated three times for 30 min each in propylene oxide. Propylene oxide was gradually exchanged by Epon812 and samples were polymerised for at least 48 hours at 60°C. Sectioning was performed with an ultramicrotome EM UC7 (Leica) and ultra-thin sections (60-100 nm) for transmission electron microscopy (TEM) or semi-thin sections for tomography (200-250 nm) were post-stained 20 min in 2% uranyl acetate in ethanol and 10 min in Reynold's lead citrate (Reynolds, 1963) to increase contrast. Additionally, grids holding sections for tomography were carbon coated. The prepared sections were analysed using a 200 kV JEM-2100 (Jeol, Munich, Germany) transmission electron microscope equipped with a TemCam F416 4k x 4k camera (Tietz Video and Imaging Processing Systems, Gauting, Germany). Electron micrographs of standard sections were taken with the EMMenu v 4.0.9.31 software (Tietz Video and Imaging Processing Systems) and further analysed using ImageJ 1.49v. The TrakEM2 plugin was used to arrange tiling series and the resulting pictures were further arranged using Inkscape 0.91 and Adobe Illustrator CS4 14.0.0. The tilt series for tomogram generation were recorded from at least -65° to +65° with 1° increments with SerialEM (Mastronarde, 2005). The picture series were further processed with eTOMO, a comprehensive toolset within the software package IMOD v4.7 (Kremer et al., 1996).

3.11.3 High pressure freezing of *C. floridanus* ovaries for TEM

The most commonly used protocols for electron microscopy involve chemical fixation of tissues. Although chemical fixation crosslinks organic molecules and preserves structures for microscopy, it most often produces artefacts as well. In order to reduce these artefacts, samples can be frozen and the frozen water can be substituted by any fixative. It is important to avoid ice crystal formation, which would otherwise destroy the tissue. High pressure freezing (HPF) allows preservation of tissues up to a thickness of 600 μm. A current of liquid nitrogen is passing through the freezing chamber and a high pressure of > 2100 bar is applied simultaneously (Weimer, 2006). Thus, the freezing process is rapid enough to take place in a few milliseconds and preserves ultra-structures in a close-to-native state (Correia and Detrich, 2009).

C. floridanus worker ovarioles were dissected in 1xPBS and held on ice during dissection. For HPF only the anterior tips (germarium) of each ovariole were taken and carefully removed from the rest. Thus, the tissues were small enough to fit into the carriers with a diameter of 3 mm and a depth of 200 μm. To avoid ice formation, 20% PVP (polyvinylpyrrolidone) in 1xPBS was applied as a freezing solution. The freezing solution was pipetted into the deep side of a carrier and single tissue sections were transferred into the solution. It is necessary to carefully remove any air bubble from the solution within the carrier to allow efficient HPF of the whole sample. Shortly before HPF, the carrier was closed with the flat side of another carrier. The resulting freezing chamber was placed inside sample holders which were closed and fixed to the aperture for HPF. The freezing process was started by loading the sample inside the high pressure freezer Leica EM HPM100 (Leica, Wetzlar, Germany). After freezing, the sample holder was released in a liquid nitrogen filled container. Samples within the freezing chambers were transferred into an unloading station and removed from the holders. During the whole process, samples were kept in liquid nitrogen and therefore at temperatures below -90 °C in order to prevent a subsequent ice crystal formation.

Freeze substitution was performed in a respective freeze substitution system EM AFS2 (Leica). At first, frozen samples were transferred into the cooled 0.1% $KMnO_4$ in acetone solution for an initial washing step. The washing solution was exchanged after one hour and samples were incubated in fresh solution for 80 h at -90°C. Afterwards, temperature was risen to -45°C and $KMnO_4$ was continuously removed by 4-5 washing steps with 100% acetone over 3 h. Subsequently, acetone was substituted by ethanol through an acetone-ethanol series with increasing ethanol concentrations. Complete removal of acetone is necessary to avoid impediment of polymerisation of the LR-White resin (Medium Grade Acrylic Resin, London Resin Company Ltd). At last, temperature was risen to 4°C for the following embedding steps. Prior to embedding, ethanol was continuously substituted by LR-White. For the embedding, gelatine capsules were filled to one-third with 100% LR-White and samples were transferred into the capsules. After careful positioning of the samples, capsules were filled completely with LR-White, closed and incubated at 50°C for 3 days. Sectioning of polymerised

samples was performed with an ultramicrotome EM UC7 (Leica) and ultra-thin sections (60-100 nm) for TEM were prepared. Post-embedding staining to increase contrast of the samples was performed as described before with 2% uranyl acetate in dH_2O and 10 min in Reynold's lead citrate (Reynolds, 1963).

3.12   Analysis of gene expression in *C. floridanus* ovarian tissue

In order to observe the expression of putative immune genes in *C. floridanus* ovarian tissue, total RNA was isolated and the corresponding cDNA generated. The gene expression in comparison to other tissues was analysed via qRT-PCR. Briefly, total RNA was isolated separately from dissected ovaries or from midguts as symbiont bearing tissue and from symbiont-free residual body parts. To obtain comparable amounts of tissue, 10 ovaries, 15 midguts and 3 residual bodies were dissected on ice in sterile, RNase-free 1xPBS and pooled. Total RNA was extracted using TRIzol® Reagent (Sigma-Aldrich) and purified through RNeasy mini kit columns (Qiagen, Hilden, Germany) including an on-column DNase digestion (RNase-Free DNase Set, Qiagen) (chapter 3.3). RNA quality was determined with a Nanodrop® (Thermo Scientific, Wilmington, DE, USA) spectrophotometer as described above. For each sample cDNA was produced by reverse transcription from 1 µg of total RNA using the RevertAid First Strand cDNA Synthesis Kit (Fermentas). Resulting cDNA was diluted to a final concentration of 10 ng/µl. For gene expression studies oligonucleotides (Table 5) were designed with Primer3 v. 0.4.0 (Koressaar and Remm, 2007; Untergasser et al., 2012). The qRT-PCR experiments were performed on a StepOnePlus™ Real-Time PCR System (Applied Biosystems, Life Technologies GmbH, Darmstadt, Germany) with PCR samples containing 1x Absolute™ PerfeCTa™ SYBR® Green FastMix™ (Rox) (Quanta Biosciences, Gaithersburg, MD, USA), gene-specific oligonucleotides (250 nM each), 1 µl of the cDNA and ddH_2O to a final volume of 20 µl. The initial enzyme activation for 5 min at 95°C was followed by 45 cycles of 5 sec denaturation at 95°C, 10 sec of annealing at 56°C and 20 sec of extension at 60°C. Fragment specificity was checked in melting curves and each biological sample was run in triplicate in the qRT-PCR.

As described above (chapter 3.5.3.3) the suitability of four putative reference genes *ribosomal protein L32* (*rpL32*), *rpL18*, *elongation factor 1-alpha* (*EF1α*) and *glyceraldehyde-3-phosphate dehydrogenase* (*GAPDH*) was determined on the basis of their expression levels within the different tissues (ovary, midgut and residual body) using the BestKeeper analysis (Pfaffl et al., 2004). Afterwards, the qRT-PCR results were averaged and relative transcription levels were calculated by the ddCt method (Livak and Schmittgen, 2001). Statistical analyses of expression data were performed using Statistica 64 v12. The data set was tested for normality between samples with the Shapiro-Wilk W Test. The average gene expression levels relative to the constitutively expressed *rpL32* (dCt = Ct(target) – Ct(*rpL32*)) of six independent samples (six biological replicates as groups of workers taken from orphaned sub-colonies either originated from colony C90 or C152) were determined for the

three different tissues. Expression levels were tested with an ANOVA in order to investigate the significance of two influencing effects, tissue and colony. "Tissue" was included as fixed factor in the analyses and nested within colony. Since workers from sub-colonies deriving from two different original colonies were analysed, "colony" was included as random factor. An LSD post hoc test was performed to investigate significant differences of dCt-values between samples.

4 Results

4.1 The *C. floridanus* immune transcriptome

4.1.1 Functional re-annotation of the *C. floridanus* genome

For the functional re-annotation of the *C. floridanus* genome, RNA of workers and larvae with or without preceding immune challenge was isolated and samples were applied separately for Illumina RNA sequencing. The obtained sequencing results provided 125,873,897 reads for the sample with immune challenge and 118,142,837 reads for the sample without treatment. The obtained Illumina sequencing reads and data from *C. floridanus* expressed sequence tags were used for the re-annotation with a GHMM based optimised Augustus software, leading to an improvement in the annotation of the previously published *C. floridanus* genome. In total 15,631 genes coding for 18,369 proteins were predicted in the new genome annotation (http://camponotus.bioapps.biozentrum.uni-wuerzburg.de/) compared to 17,064 proteins in the previous annotation which is labelled cflo_OGSv3.3 (http://hymenopteragenome.org/camponotus/?q=genome_consortium_datasets). The respective sequences of predicted proteins are highly similar to protein sequences of other ant species which have been genome sequenced recently.

All proteins within the NCBI database were analysed using the BLAST tool and best sequence-similar hits of each *C. floridanus* sequence were generated with an e-value cut-off of $1e^{-5}$ (analysis performed by Shishir K. Gupta). The protein sequences reflect the phylogenetic relations, and *C. floridanus* protein sequences are most closely related to protein sequences of *Cerapachys biroi* (29.2% similarity according to best sequence hits), *Acromyrmex echinatior* (26.81%), *Solenopsis invicta* (14.98%) and *Harpegnathos saltator* (11.12%). Further, best sequence-similarity hits were distributed among other insects such as *Bombus* and *Apis* species, *Nasonia vitripennis*, *Tribolium castaneum*, *Acyrthosiphon pisum*, *Drosophila melanogaster* and others (Figure 7).

The optimised exon level accuracy of 80.8% allowed an improved analysis of splicing events. The new genome annotation predicted 1,928 genes with two or more alternative transcripts resulting in an overall number of 4,666 alternative transcripts. In contrast, the previous annotation v3.3 revealed 2,538 genes coding for 7,583 alternative transcripts.
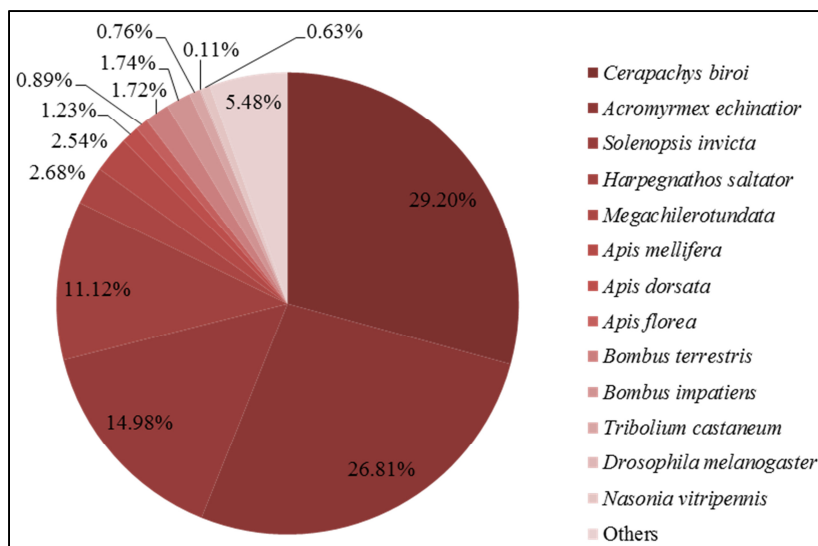
Figure 7 Distribution of closest related sequences to *C. floridanus* proteins. Sequences of *C. floridanus* proteins were compared with all proteins within the NCBI database. The best sequence-similar hits are distributed over several insect species with *C. floridanus* protein sequences (of all assigned best sequence-similar hits) being closest related to *Cerapachys biroi* proteins (29.2% best hits).

Furthermore, the software Blast2GO was used to assign level-2 gene ontology (GO) terms out of the three categories "biological process", "molecular function" and "cellular component" to 8,490 out of the 18,369 predicted proteins in the new annotation (performed by Dr Frank Förster and Shishir K. Gupta). Within the category "biological process" (Figure 8A), most of the proteins were assigned to three major terms: (i) "cellular process" (19.96%), (ii) "metabolic process" (16.72%), and (iii) "single-organism process" (16.60%). While another major set of proteins was assigned to the GO term "biological regulation" (8.96%), the categories "response to stimulus" (6.38%), "localisation" (5.71%), "multicellular organismal process" (5.51%), "signalling" (5.12%), "developmental process" (5.06%) and "biogenesis" (4.62%) comprised protein sets of about the same size. Interestingly, only 0.53% of the proteins were categorised into the level-2 GO term "immune system process". Regarding the category "molecular function" (Figure 8B), proteins with "catalytic function" (40.07%) and involved in "binding" (39.09%) were most abundant. The remaining proteins (20.84%) were assigned to the other categories, e.g. "enzymatic regulator activity", "transporter activity", or "receptor activity". With regard to the level-2 GO category "cellular compartment" (Figure 8C), the largest fractions among the proteins were allocated to the subcategories "cell" (33.74%), "organelle" (22.97%), "membrane" (18.82%) and "macromolecular complex" (14.66%). In comparison to the previous genome annotation, the number of proteins successfully assigned to GO terms increased from 18% (7,143 proteins of the genome annotation cflo_OGSv3.3) to 23% of newly annotated proteins. Therefore, the transcriptome data permitted the refinement of the functional annotation of *C. floridanus* proteins.
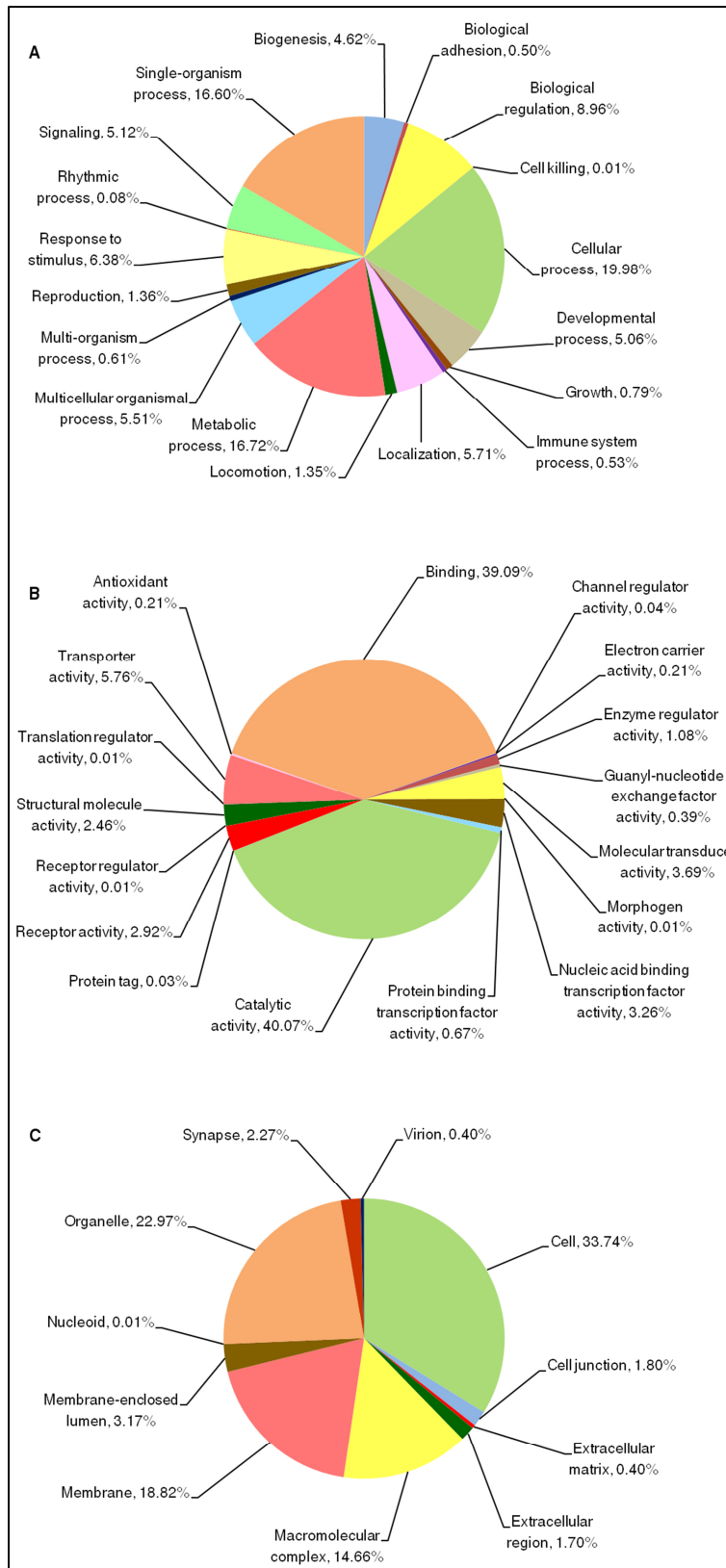
Figure 8 Distribution of functions in *C. floridanus* proteins. 8,490 proteins were categorised in level-2 GO terms for **(A)** "biological process", **(B)** "molecular function", and **(C)** "cellular component". (Figure generated by Shishir K. Gupta; Gupta et al. (2015))

4.1.2   The immune signalling pathways of *C. floridanus*

As mentioned above, 0.53% of *C. floridanus* proteins were assigned to the subcategory "immune system process". Furthermore, an orthology analysis of immune-related proteins was performed by Shishir K. Gupta. For this purpose, protein sequences of the ant were compared to those of *D. melanogaster*, the honey bee *A. mellifera* and the parasitoid wasp *Nasonia vitripennis*. The high sequence similarity of two proteins makes them orthologs allowing the assumption that their functions and modes of action are conserved. As shown in Figure 9, *C. floridanus* shares most orthologs with *N. vitripennis* (307), 271 orthologs with *D. melanogaster*, and 221 orthologs with *A. mellifera*. Interestingly, the 65 protein orthologs shared by all three species mostly belong to the core signalling pathways of the immune system, the Toll, the IMD, the Jak-Stat, and the JNK pathway, but they also comprise proteins involved in microbial recognition as well as serine proteases (Appendix II, Table 20). Furthermore, genes coding for 510 proteins or isoforms of proteins with immune-related function were identified in *C. floridanus* and recently published (Additional file 7: Table S5 published in Gupta et al. (2015)). These proteins were predicted to be involved in recognition and immune signalling, in phagocytosis, melanisation, and coagulation, and they also comprise AMPs. In the following the major immune signalling pathways in *C. floridanus* will be described.
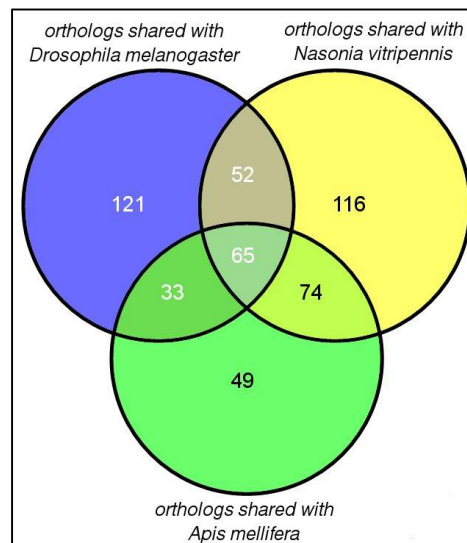
Figure 9 Immune-related genes of *C. floridanus* shared with three other insect species. The Venn diagram shows *C. floridanus* protein orthologs shared with the fruit fly *D. melanogaster*, the honey bee *A. mellifera* and the wasp *N. vitripennis*. 65 protein orthologs are shared by all three species. (Figure generated by Shishir K. Gupta; Gupta et al. (2015))

Signalling pathways are the interconnecting parts between the recognition of microbes and the production of immune effectors. In general, pattern recognition receptors (PRRs) recognise microbial foreign structures, the so-called MAMPs, e.g. bacterial peptidoglycan (PGN) and β-1,3-glucans. Upon binding of the receptors to their ligands, intracellular signalling cascades are stimulated and transcription factors are transported into the nucleus, where immune gene expression is activated. Such genes encode for example various AMPs. The most well understood immune signalling pathways are those of the fruit fly *D. melanogaster* (Lemaitre and Hoffmann, 2007). Three databases

(KEGG, Fly Reactome and INOH) as well as literature analysis were used to reconstruct the immune signalling pathways of *D. melanogaster* which were then used as a matrix for the mapping of the respective *C. floridanus* proteins based on sequence homologies (performed by Shishir K. Gupta). In general, the immune-related signalling repertoire of *C. floridanus* is conserved and presents high similarities to the immune components identified by experiments performed in the fruit fly (Aggarwal and Silverman, 2008; Hoffmann, 2003; Lemaitre and Hoffmann, 2007).

The Toll signalling pathway mostly responds to infections with fungi or Gram-positive bacteria and is necessary for the production of antimicrobial effectors as well as for haemolymph proliferation. Therefore, the Toll pathway coordinates both, the cellular and the humoral immune response (Qiu et al., 1998; Valanne et al., 2011). Lysine-type PGN and β-1,3-glucans are recognised by extracellular PRRs, and the Spätzle processing enzyme (SPE) activates the protein Spätzle via a proteolytic cleavage cascade. Spätzle can then bind to the membrane bound Toll receptor protein, and an intracellular signalling cascade is initiated upon dimerization of the receptor. As a result, the relish transcription factors Dorsal and the DIF (Dorsal-related immunity factor) are transported into the nucleus and the expression of antimicrobial effectors is induced (Valanne et al., 2011).

Three PRRs channelling signals from immune stimuli into the Toll signalling pathway were found to be encoded by the *C. floridanus* genome. Based on sequence homology, the peptidoglycan recognition protein PGRP-SA (Cflo_N_g8526t1) is probably recognising lysine-type PGN of Gram-positive bacteria as it has been shown for *Drosophila* PGRP-SA (Aggarwal and Silverman, 2008). It was speculated that the *Drosophila* PGRP-SA further discriminates between DAP-type and lysine-type PGN by specifically hydrolysing the peptide bond connecting DAP and D-alanine thereby preventing the activation of Toll signalling by Gram-negative DAP-type PGN (Chang et al., 2004).

Two proteins were annotated as beta-1,3-glucan-binding proteins or Gram-negative binding proteins (GNBPs; Cflo_N_g15215t1 and Cflo_N_g5742t1). Both proteins show high homologies to both *D. melanogaster* proteins, GNBP1 and GNBP3. It was shown that the *Drosophila* GNBP1 recognises lysine-type PGN as well, whereas GNBP3 recognises fungal beta-glucans (Gottar et al., 2006). However, previous data gained from infection experiments in *C. floridanus* revealed the long-lasting upregulation of only one GNBP (Cflo_N_g5742t1, previously annotated as EFN66519.1) after injection of the Gram-positive *Micrococcus flavus* (Ratzka et al., 2011). Therefore, the GNBP (Cflo_N_g5742t1) may be the PRR recognising lysine-type PGN (Figure 10). Additionally, the *C. floridanus* genome encodes the serine protease Persephone (Psh) and its negative regulatory protein Necrotic (NEC). In *Drosophila*, Psh is known to recognise entomopathogenic moulds independently from the GNBP3, albeit not relying on recognition of cell wall structures. The Psh-dependent pathway is probably activated by the recognition of virulence factors (proteases and chitinases) such as PR1A produced and released by the mould *Metarhizhium anisopliae* to break through the host cuticle (Aggarwal and Silverman, 2008; Gottar et al., 2006). Additionally, it was shown that the *Drosophila*

Psh-mediated so-called "danger signal" pathway is also activated by proteases secreted by Gram-positive bacteria (Ashok, 2009; El Chamy et al., 2008). Thus, recognition of bacterial or fungal virulence factors via Persephone (Cflo_N_g8442t1) might lead to activation of the serine protease cascade through the SP Spirit in *C. floridanus* (Figure 10).
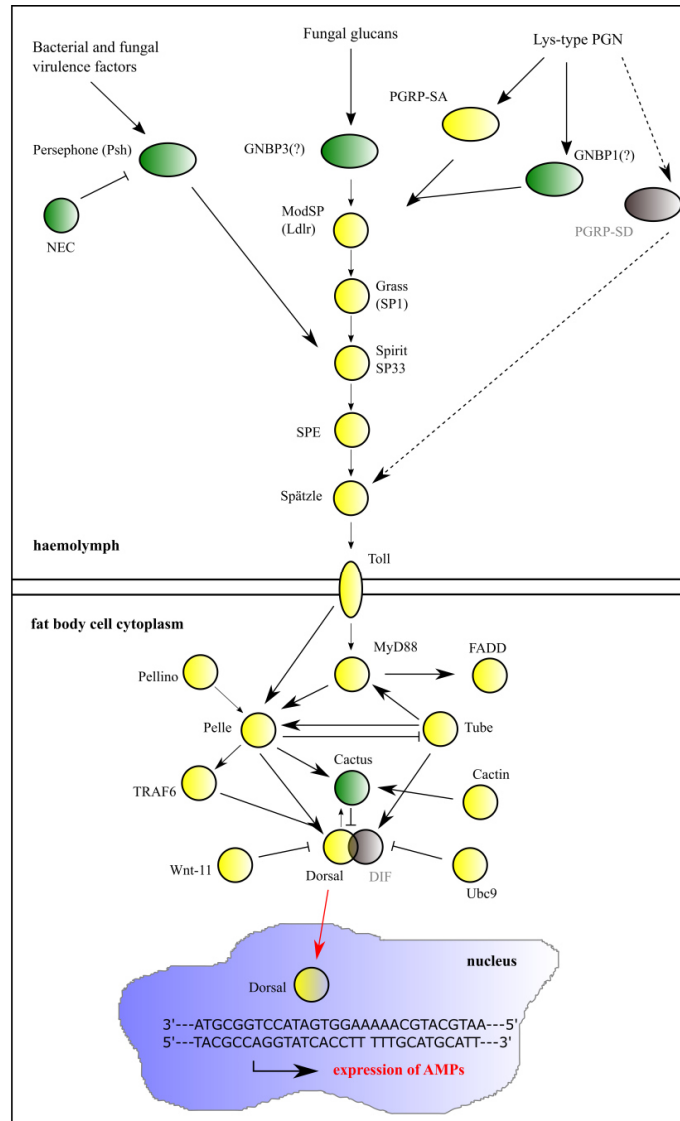

Figure 10 The Toll signalling pathway in *C. floridanus*. Based on sequence homology, signalling components were identified and mapped on the immune network of *D. melanogaster*. Hence, the names of *C. floridanus* proteins correspond to the respective *Drosophila* proteins. Connections with a pointed arrowhead indicate positive regulation, whereas connections ending in a bar indicate negative regulation. Components which showed significant upregulation upon immune challenge are highlighted in green. Proteins presented in grey colour were not identified in the *C. floridanus* genome. (Figure based on Gupta et al. (2015))

Activation of the Toll pathway through either one of the recognition receptors leads to initiation of a serine protease cascade. The protease activity of Spirit is triggered via Psh or Grass, which receive signals from bacterial and fungal pathogens. Finally, protein Spätzle is activated upon cleavage by SPE and binds to the membrane-bound receptor Toll inducing its dimerization (Aggarwal and Silverman, 2008; Valanne et al., 2011). Homologs for both proteins, Spätzle (Cflo_N_g12735t1) and Toll (Cflo_N_g8278t1), as well as for most of the intracellularly active proteins of the immune

signalling cascade were predicted in the *C. floridanus* genome. Intracellularly, a complex consisting of the Myeloid differentiation primary response protein MyD88 (Cflo_N_g11593t1) and Tube (Cflo_N_g3789t1) is recruited upon Toll dimerization, and itself recruits the kinase Pelle (Cflo_N_g1330t1). In *Drosophila*, the kinase Pelle is activated and leads either directly or indirectly to the phosphorylation and ubiquitin/proteasome-mediated degradation of the IκB homolog Cactus (Fernandez et al., 2001). Degradation of the latter results in the release of the NF-κB transcription factors Dorsal and DIF. Both factors are translocated into the nucleus, and the expression of immune responsive genes, e.g. coding for AMPs, is induced (Aggarwal and Silverman, 2008). In the *C. floridanus* genome, a gene encoding the NF-κB inhibitor Cactus (Cflo_N_g14414t1) as well as a gene coding for the transcription factor Dorsal (Cflo_N_g3305t1) were found. The absence of a DIF encoding gene suggests that only the unique NF-κB transcription factor Dorsal is required for Toll signalling mediated induction of AMP expression in the ant.

IMD signalling induced AMP expression is mostly triggered after infections with Gram-negative bacteria (Figure 11) (Lemaitre and Hoffmann, 2007; Myllymaki et al., 2014). In the original annotation of the *C. floridanus* genome, only one PRR recognising DAP-type PGN and feeding into the IMD signalling pathway was identified and annotated as PGRP-LE (GenBank Acc. No.: EFN63542). However, the *C. floridanus* genome revealed that the predicted protein shows an overall higher homology to the *D. melanogaster* or *A. mellifera* protein PGRP-LC. Thus, the *C. floridanus* protein was re-annotated as PGRP-LC (Cflo_N_g10272t1), and it appears to be the single PRR in the ant which recognises DAP-type PGN. In *Drosophila*, the membrane integrated PGRP-LC recognises extracellular PGN, whereas the soluble PGRP-LE is thought to trigger IMD pathway activation by recognition of extracellular and intracellular PGN (Kaneko et al., 2006; Myllymaki et al., 2014). However, no homolog for the *Drosophila* PGRP-LE protein was found in the *C. floridanus* genome. In *Drosophila*, PGN-mediated activation of PGRP-LC results in the recruitment of a complex consisting of the death domain protein Imd, the caspase-8 homolog Dredd and the adaptor protein FADD. The E3-ligase inhibitor of apoptosis 2 (IAP2) is involved in activation of Dredd via ubiquitination. Furthermore, the *Drosophila* IAP2 ubiquitinates Imd after the activated Dredd has cleaved Imd. A complex consisting of the Mitogen-activated protein kinase MAPKKK7 (TAK1) and the Tab2 protein is recruited and activated resulting in phosphorylation of the *Drosophila* IKK complex. Finally, the transcription factor relish is activated through phosphorylation of several serine residues. The protein's N-terminal part is translocated into the nucleus, and transcription of genes coding for various AMPs, such as diptericin and cecropin, is initiated. Based on sequence homology analyses, most components of the IMD pathway were identified in *C. floridanus* (Figure 11). However, the signalling cascade is not only lacking the gene encoding PGRP-LE but also genes encoding Kenny and Caudal. In *Drosophila*, the protein Kenny is the regulatory subunit of the IKK complex. Whereas *D. melanogaster* mutants lacking Kenny are highly susceptible to bacterial infections, the Kenny subunit was also not identified in the genomes of *A. mellifera*, *N. vitripennis* and other ant species.

This may suggest a common characteristic of the hymenopteran IKK complex, but it is not known whether the lack of the Kenny subunit has any influence on the function of the complex.
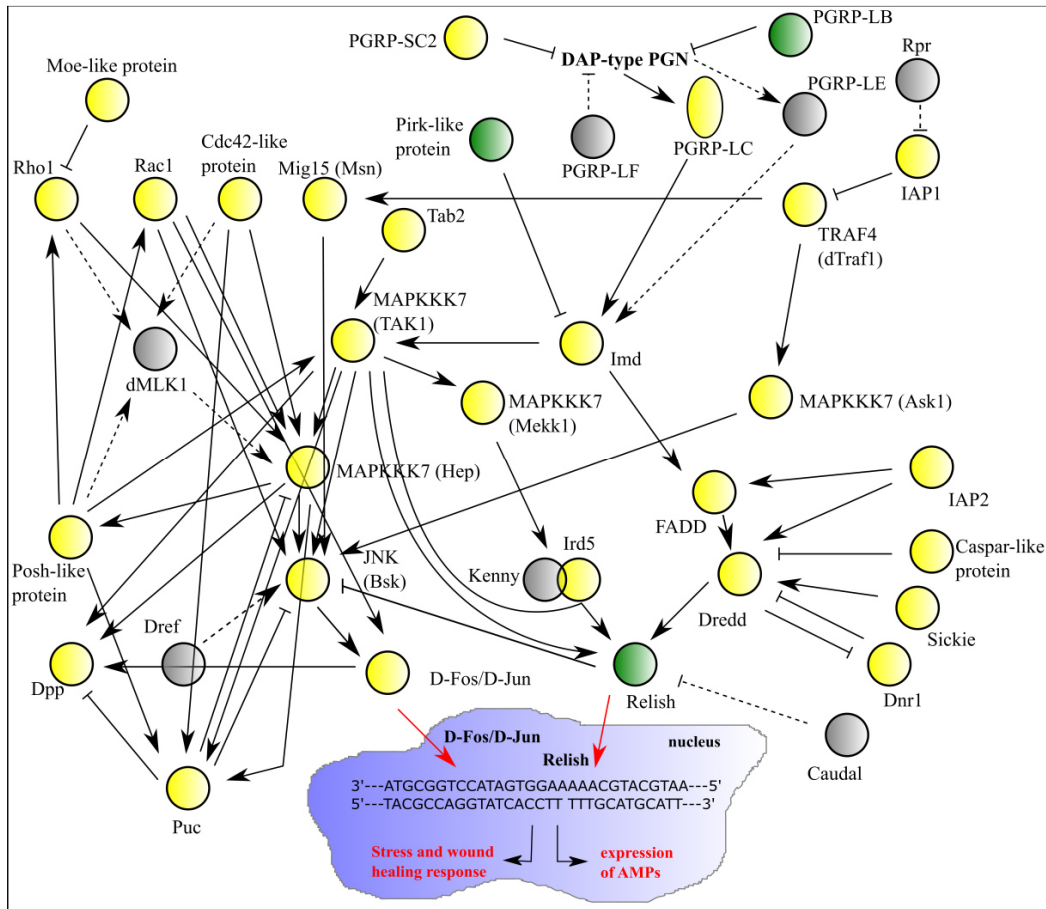
Figure 11 The IMD and JNK signalling pathway in *C. floridanus*. Based on sequence homology, signalling components were identified and mapped on the immune network of *D. melanogaster*. Therefore, the names of the *C. floridanus* proteins correspond to the respective *Drosophila* proteins. Connections with a pointed arrowhead indicate positive regulation, whereas connections ending in a bar indicate negative regulation. Components which showed significant upon immune challenge are highlighted in green. Proteins presented in grey colour were not identified in the *C. floridanus* genome. The oval shape of PGRP-LC marks it as a membrane-integrated protein. (Figure based on Gupta et al. (2015))

Furthermore, it was shown that the intestinal homeobox gene coding for Caudal is expressed in the *Drosophila* midgut. It was suggested that tissue homeostasis is maintained by the expression of Caudal and other negative regulators of the IMD pathway in order to prevent an extensive IMD signalling mediated immune response towards commensal microbes (Myllymaki et al., 2014; Ryu et al., 2008). The regulatory mechanisms of the IMD signalling pathway may also serve in control and tolerance of the *C. floridanus* endosymbiont *B. floridanus*. Two PRRs with putative regulatory function were identified in the ants' genome: PGRP-LB (Cflo_N_g103t1) and PGRP-SC2 (Cflo_N_g102t2). Both PRRs contain an amidase domain. Therefore, the proteins may be involved in down-modulation of the IMD signal transduction pathway by cleaving the DAP-type PGN into non-immunostimulating molecules (Bischoff et al., 2006; Kurata, 2014; Royet et al., 2011). Recently, it was shown that PGRP-LB (Cflo_N_g103t1) was highly upregulated only in the midgut tissue of *C. floridanus* during pupation. During this phase, the bacterial endosymbionts massively multiply which may be supported

by the down-modulation of an IMD signalling mediated immune response in the midgut tissue of the ants (Ratzka et al., 2013a).

The IMD signalling pathway in *C. floridanus* may be highly interactive with the JNK signalling pathway as it has been shown for *D. melanogaster*. The Transforming growth factor β-activated kinase 1 (TAK1) acts as a JNK kinase and activates JNK signalling in addition to the phosphorylation of the IKK complex. The *Drosophila* protein Hemipterous (Hep) is activated and then phosphorylates the protein Basket (Bsk/ dJNK) resulting in the overall activation of transcription factors (e.g. F-Dos/ D-Jun) (Aggarwal and Silverman, 2008; Boutros et al., 2002; Hetru and Hoffmann, 2009). Almost every component of the JNK signalling pathway was identified in the *C. floridanus* genome. Therefore, it can be assumed that the ants can initiate the expression of genes relevant for wound healing, inflammatory and cellular immune responses upon JNK activation, as it has been shown for *Drosophila* (Ip and Davis, 1998; Rämet et al., 2002a; Sluss et al., 1996).

The Jak-Stat pathway was originally described due to its involvement in embryonic segmentation in *Drosophila* (Binari and Perrimon, 1994). The major components of this pathway comprise the ligands, Unpaired (Upd), of which three could be identified in *Drosophila* (Upd, Upd2 and Upd3), the receptor Domeless (Dome), the downstream Janus kinase Hopscotch (Hop) and the transcription factor Stat92E. Previous works indicate that the Upd ligands induce phosphorylation and activate the pathway through the transmembrane receptor Dome. The downstream signalling is induced via Hop and Stat92E with the latter containing a DNA-binding domain and a C-terminal tyrosine residue. Thus, the transcription factor can be activated upon phosphorylation and is translocated into the nucleus (Agaisse and Perrimon, 2004; Brown et al., 2001; Chen et al., 2002; Harrison et al., 1998). On the basis of sequence homologies, almost all components of the Jak/Stat pathway could be identified in *C. floridanus* (Figure 12). Surprisingly, no orthologs for the ligands Upd, Upd2, and Upd3, which channel signals into the pathway, were identified. However, these components are also absent from the genomes of other insects, for example the honey bee (Evans et al., 2006). Furthermore, only one set of downstream effector proteins was identified in the ant. Similar to *A. mellifera*, no homologs of the Turandot proteins were found, but several thioester-containing proteins (TEPs) were identified (Evans et al., 2006). TEPs show significant sequence similarities to vertebrate complement factors and the α_2-macroglobulin family of protease inhibitors. Most TEPs contain a common motif (CGEQ) which defines the thioester site. Through this site, a covalent bond to the microbial surface is formed. Thereby, TEPs supposedly act as opsonins promoting phagocytosis (Bou Aoun et al., 2011). While it has been shown that TEP1 is involved in phagocytosis in *Anopheles gambiae* and several TEPs contribute to phagocytosis in *Drosophila* S2 cells (Blandin et al., 2004; Stroschein-Stevenson et al., 2006), no qualitative defect in haemocyte mediated phagocytosis was observed in *Tep* mutant *D. melanogaster* which were infected with bacteria (Bou Aoun et al., 2011). Three TEPs are present in the *C. floridanus* genome and were further analysed by Shishir K. Gupta: TEP1 (Cflo_N_g4492t1),

TEP2 (Cflo_N_g7345t1), and TEP3 with two alternative transcripts (Cflo_N_g9745t1 and Cflo_N_g9745t2). Only TEP1 and TEP2 contain the CGEQ motif in the thioester site and are therefore believed to promote binding to microbial surfaces. Additionally, the expression of TEP2 was upregulated in immune challenged *C. floridanus*. Though the amino acid sequence of the TEP3 protein lacks the CGEQ motif, it may be relevant for the complement system in *C. floridanus*. It was proposed that TEPs without the defining motif in the thioester site may act as adaptors in the formation of the membrane attack complex similar to vertebrate complement factors (Blandin and Levashina, 2004).

Figure 12 The Jak-Stat signalling pathway in *C. floridanus*. Based on sequence homology, signalling components were identified and mapped on the immune network of *D. melanogaster*. Hence, the names of *C. floridanus* proteins correspond to the respective *Drosophila* proteins. Connections with a pointed arrowhead indicate positive regulation, whereas connections ending in a bar indicate negative regulation. Proteins presented in grey colour were not identified in the *C. floridanus* genome. (Figure based on Gupta et al. (2015))

4.1.3 The *C. floridanus* AMP repertoire in comparison to other insect species

The potential of an animal's immune system may be especially represented in the extent of its set of antimicrobial effectors such as antimicrobial peptides (AMPs). Addressing this important matter, the AMP repertoires of several insect species have been analysed. The analysis of the *C. floridanus* transcriptome data revealed the existence of only a low number of genes encoding known AMPs, such

as the genes coding for hymenoptaecin and two defensins which have already been identified and characterised in a previous work (Ratzka et al., 2012a). Also, two isoforms of a tachystatin-like peptide (U8-agatoxin-Ao1a-like; Cflo_N_g5938t1 and Cflo_N_g5938t2 in the newly annotated genome) were identified confirming the results of another previous work (Zhang and Zhu, 2012). Furthermore, two genes encoding crustin-like peptides (Cflo_N_g4955t1 and Cflo_N_g4956t1) were identified. Crustins from crustaceans are known to exhibit antimicrobial activity and contain an N-terminal signal sequence as well as at least one whey acidic protein (WAP) domain at the C-terminus (Smith et al., 2008; Zhang and Zhu, 2012). Most crustins are cationic cysteine-rich AMPs with molecular weights of 7-14 kDa and the WAP domain is located at their C-terminus. This domain exhibits eight cysteine residues in a conserved arrangement which is as follows:

$$C_1 - (X_n) - C_2 - (X_n) - C_3 - (X_5) - C_4 - (X_5) - C_5 - C_6 - (X_{3/5}) - C_7 - (X_{3/4}) - C_8.$$
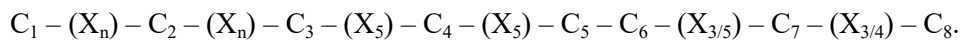
"X" indicates any amino acid and "X_n" refers to a stretch of n amino acids. The central four cysteines C_3-C_6 form the basis of the four-disulfide core of the WAP domain (Smith et al., 2008). The WAP four-disulfide core domain protein 2-like isoform X1 (Cflo_N_g4955t1; previously annotated as waprin Thr1-like isoform X1, XP_011258021.1) contains one WAP domain but lacks the aromatic amino acids-rich region which is typical for the Hymenoptera type IV crustins (Zhang and Zhu, 2012). Interestingly, the *C. floridanus* crustin-like peptide now annotated as waprin-Phi1 (Cflo_N_g4956t1; previously annotated as waprin Phi1-like isoform X1, XP_011258020.1) also lacks the aromatic amino acids-rich region but contains two WAP domains (Figure 13). Therefore, the waprin-Phi1 was categorised as a member of the double WAP domain (DWD) protein family (Gupta, 2016). Previously, it was shown that DWD proteins are involved in the regulation of proteinase activity in shrimp (Du et al., 2009; Suthianthong et al., 2011), whereas a DWD protein from the Chinese mitten crab *Eriocheir sinensis* exhibits antimicrobial activity against the Gram-negative *Vibrio anguillarum* and yeast (Li et al., 2012). Interestingly, the previously described *C. floridanus* crustin (Zhang and Zhu, 2012) was now identified as a part of the C-terminus of the C3 and PZP-like alpha-2-macroglobulin domain-containing protein 8 (Cflo_N_g8245t1) in the re-annotated genome of *C. floridanus* (Figure 13).
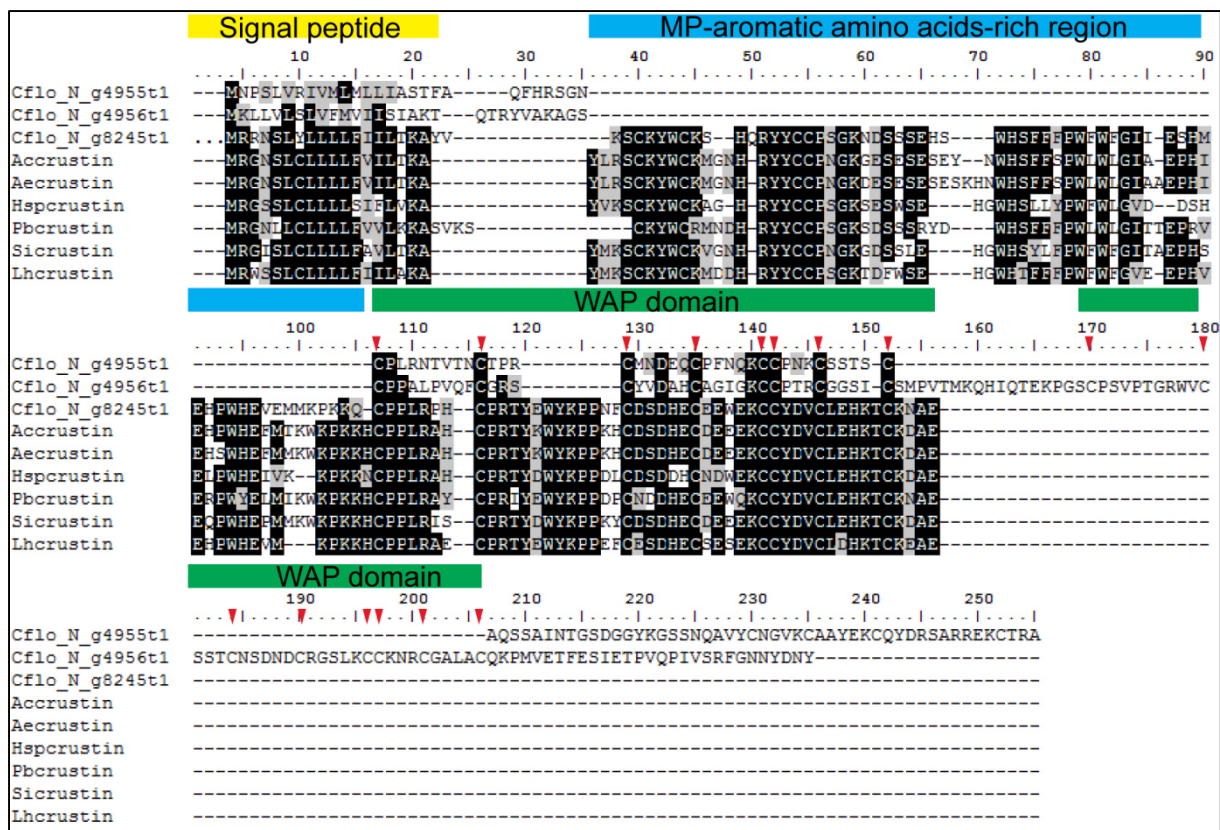
Figure 13 Alignment of crustin-like peptides of *C. floridanus* and six other ant species. The amino acid sequences of two *C. floridanus* peptides which contain either one C-terminal WAP domain (Cflo_N_g4955t1) or two WAP domains (Cflo_N_g4956t1) were aligned with the C-terminal sequence of the C3 and PZP-like alpha-2-macroglobulin domain-containing protein 8 (Cflo_N_g8245t1) and crustin sequences from *Atta cephalotes* (Ac), *Acromyrmex echinatior* (Ae), *Harpegnathos saltator* (Hs) *Pogonomyrmex barbatus* (Pb), *Solenopsis invicta* (Si) and *Linepithema humile* (Lh). Crustin-like peptides generally contain a signal peptide (yellow) and the C-terminal WAP domain (green). The previously identified crustins from six different ant species additionally exhibit an MP-aromatic amino acids-rich region (blue) (Zhang and Zhu, 2012) and the previously identified crustin of *C. floridanus* is now a part of the C-terminus of the C3 and PZP-like alpha-2-macroglobulin domain-containing protein 8 (Cflo_N_g8245t1) which is indicated by three dots at the beginning of the sequence. Identical amino acid residues are highlighted with a black background and similar ones with a grey background. The eight cysteine residues which are characteristic for a WAP domain are marked with red arrowheads.

With the exception of the Crustin-like peptide, which was not identified in the genome of the clonal raider ant *Cerapachys biroi*, homologs of genes encoding the AMPs of *C. floridanus* were found in the sequenced genomes of other ants (Oxley et al., 2014; Zhang and Zhu, 2012). Additionally, the genomes of the ant species *Harpegnathos saltator*, *Pogonomyrmex barbatus*, *Atta cephalotes*, *Achromyrmex echinatior* and *C. biroi* encode the AMP abaecin (Table 9). Interestingly, the overall number of antimicrobial peptide genes in the ants and in the social honey bee *A. mellifera* appears to be quite low when compared to the solitary wasp *Nasonia vitripennis* (Hymenoptera), for which 44 AMPs were identified (Tian et al., 2010).

Table 9 List of antimicrobial peptides of *C. floridanus* and other ant species. The information presented in the table was gained by analysis of several genomes by the use of HMMER3 applied to the AMPer database (Fjell et al., 2007), the sequence analyses with BlastP and tBlastn, and literature analysis performed by Shishir K. Gupta (Bonasio et al., 2010; Nygaard et al., 2011; Oxley et al., 2014; Sackton et al., 2013; Smith et al., 2011a; Smith et al., 2011b; Suen et al., 2011; Tian et al., 2010; Wurm et al., 2011; Zhang and Zhu, 2012). (Table based on Gupta et al. (2015))

AMPs	*C. floridanus*	*H. saltator*	*L. humile*	*P. barbatus*	*A. cephalotes*	*S. invicta*	*A. echinatior*	*C. biroi*	*A. mellifera*	*N. vitripennis*
Hymenoptaecin	1	2	1	1	1	1	3	1	1	2
Defensin	2	2	1	5	1	2	1	1	3	5
Tachystatin-like	2	2	3	2	1	3	2	3	1	3
Crustin-like	2*	1	1	1	1	1	1	-	-	1
Abaecin	-	1	-	1	1	-	1	1	1	-
Melittin	-	-	-	-	-	-	-	-	1	-
Apisimin	-	-	-	-	-	-	-	-	1	-
Apidaecin	-	-	-	-	-	-	-	-	5	-
Navitripenicin	-	-	-	-	-	-	-	-	-	4
Nasonin	-	-	-	-	-	-	-	-	-	14
Nabaecin	-	-	-	-	-	-	-	-	-	4
Glynavicin	-	-	-	-	-	-	-	-	-	7
Hisnavicin	-	-	-	-	-	-	-	-	-	5
Nahelixin	-	-	-	-	-	-	-	-	-	1
Sum of AMPs	**7**	**8**	**5**	**10**	**5**	**7**	**8**	**6**	**13**	**46**

*) Please note that the previously identified *C. floridanus* crustin (Zhang and Zhu, 2012) only exhibits the C-terminus of a protein with higher molecular weight (Cflo_N_g8245t1) in the re-annotated genome of the ant. Therefore, only the two annotated waprins/ crustins (Cflo_N_g4955t1 and Cflo_N_g4956t1) were included in this table.

4.1.4 Differential expression of immune-related genes in *C. floridanus*

The *C. floridanus* genes with putative immune functions were identified mainly based on sequence homologies with a strong focus on similarities especially in functional domains. However, the transcriptome data set was further used to identify genes differentially expressed upon immune challenge. The fact that a gene is up- or downregulated in the ants after an immune challenge strongly suggests its involvement in the overall immune response but also in other stress responses.

For the transcriptome analysis samples of pooled larvae and adult workers either injected with a mixture of heat-killed *M. luteus* and *E. coli* or untreated were used. In order to identify differentially expressed genes between immune-challenged and untreated samples, two programs were used: Cufflinks (Trapnell et al., 2012) and *DESeq* (Anders and Huber, 2010). Combined data analysis, performed by Shishir K. Gupta and Frank Förster, allowed the identification of 257 differentially expressed transcripts within the immune challenged sample with > 95% likelihood (Gupta et al., 2015; Ratzka, 2012). Among the upregulated genes are those encoding PRRs such as the GNBP (Cflo_N_g5742t1) and the Hemolymph lipopolysaccharide-binding protein/tyrosine-protein kinase Btk29A (Cflo_N_g8897t1); genes encoding serine proteases, e.g. Stubble (Cflo_N_g12631t2), Persephone (Cflo_N_g8442t1), and Snake (Cflo_N_g8989t1); genes coding for proteins involved in downstream signalling and transcription induction like the nuclear factor NF-kappa-B p110 subunit Relish (Cflo_N_g6082t1); as well as genes encoding proteins involved in general stress responses, e.g. Cytochrome P450 6A1 (Cflo_N_g11706t1), Cytochrome b5 (Cflo_N_g5800t1), Transferrin (Cflo_N_g7714t1), and Apolipoprotein D/LIM domain kinase 1 (Cflo_N_g13783t1). The upregulation of the expression of stress response related genes appears to be linked to the introduction of septic injuries and has been reported previously for *C. floridanus* and other insects (Altincicek et al., 2008; Gerardo et al., 2010; Ratzka et al., 2011; Zhang et al., 2014a).

Besides the expression of the antimicrobial peptide gene *defensin-1* (JN989495.1), the expression of several negative immune regulator genes was also found to be upregulated, e.g. *serine protease inhibitor dipetalogastin* (Cflo_N_g3252t1), *PGRP-LB* (Cflo_N_g103t1), *NF-kappa-B inhibitor cactus* (Cflo_N_g14413t1). Interestingly, the expression of many genes involved in metabolism and storage such as *chymotrypsin-1* (Cflo_N_g907t1), *sugar transporter ERD6-like 6* (Cflo_N_g4569t1), *trypsin-1* (Cflo_N_g4443t1), *maltase-1* (Cflo_N_g8231t1) or *lipase-3* (Cflo_N_g13955t1) was downregulated after bacterial infection.

For validation of the transcriptome data and to gain a further and more detailed view into the regulation of gene expression towards bacterial immune challenge, the expression of several genes with putative immune function was analysed via qRT-PCR. For this purpose larvae and adult workers of *C. floridanus* were treated separately. The animals were immune challenged with a mixture of heat-killed *M. luteus* and *E. coli* via injection, and RNA was isolated from either immune-challenged or untreated animals. Putative immune genes chosen for qRT-PCR analysis are listed in Table 10.

Additional information with respect to the significance of relative gene expression levels (2^{-ddCt}-values) obtained via qRT-PCR is given in Table 21 and Table 22 (Appendix III).

Table 10 Differentially expressed genes with immune-related function. The expression of 37 selected genes after immune challenge was analysed separately in larvae and workers via qRT-PCR revealing a stage-specific gene regulation. The table shows genes significantly up and downregulated 12 h after pricking of larvae or workers with a 1:1 mix of Gram-negative and Gram-positive bacteria (mean ratio L2; mean ratio W2). The corresponding results for differential gene expression within the transcriptome data set were generated using Cufflinks and DESeq, and are given as "x-fold change". "Mean ratios L2+W2" resulting from qRT analyses were used for better comparison of expression data between the two experimental settings, qRT-PCR and Illumina sequencing. Additionally, a heat map highlights each ratio ranging from downregulation (< 0.5; dark blue) to upregulation (> 2.0; dark red). († = no values available in the data set)

Gene name	Annotation	Cflo-New	mean ratio L2	mean ratio W2	mean ratio L2+W2	Cuffdiff x-fold change	DESeq x-fold change
Housekeeping Genes							
rpL32	60S ribosomal protein L32	Cflo_N_g6999t1	1.06	1.01	1.03	0.96	0.88
Potential Immune Genes							
scav	Scavenger receptor class B member 1	Cflo_N_g9950t1	0.69	0.26	0.47	0.15	0.46
Posub	Phenoloxidase subunit A3	Cflo_N_g1918t1	0.36	0.52	0.44	0.17	0.17
MPI	Metalloproteinase inhibitor 3	Cflo_N_g622t1	5.92	0.92	3.42	2.14	1.90
yellow	Protein yellow	Cflo_N_g14239t1	2.26	2.31	2.28	2.23	2.17
SOCS2	Suppressor of cytokine signaling 2	Cflo_N_g4920t1	4.76	1.68	3.22	2.45	2.13
cact1	NF-kappa-B inhibitor cactus	Cflo_N_g14413t1	7.35	2.97	5.16	2.49	2.44
vitel	Vitellogenin-3 (Fragment)	Cflo_N_g8262t1	1.11	9.18	5.15	3.83	3.64
transf	Transferrin	Cflo_N_g7714t1	12.10	1.48	6.79	5.58	5.28
hp70940	Hypothetical protein	Cflo_N_g9484t1	5.42	8.30	6.86	1.35	1.10
PHR	Parathyroid hormone/parathyroid hormone-related peptide receptor	Cflo_N_g6985t1	35.28	6.77	21.03	13.64	12.78
ester	Esterase FE4/Carboxylic-ester hydrolase	Cflo_N_g8597t1	40.53	1.85	21.19	27.21	26.38
hp67112	Hypothetical protein	Cflo_N_g6748t1	936.46	9.44	472.95	276.28	213.29
PGRP-LB	PGRP-LB	Cflo_N_g103t1	12.07	3.14	7.60	6.33	6.48
PGRP-2	PGRP2 (=PGRP-SA)	Cflo_N_g8526t1	6.94	1.93	4.44	1.00	1.88
Rel	Nuclear factor NF-kappa-B p110 subunit	Cflo_N_g6082t1	2.57	4.51	3.54	2.90	2.53
Hym	Hymenoptaecin	Cflo_N_g14777t1	409.64	17.80	213.72	†	†
Def	Defensin-1	JN989495.1	10.37	7.34	8.86	27.12	24.15
TyrOH	Tyrosine 3-monooxygenase	Cflo_N_g5222t2	11.44	7.55	9.49	9.18	8.68
TEP2	Alpha-2-macroglobulin	Cflo_N_g7345t1	22.62	0.84	11.73	1.39	1.31
NOS1	NADPH--cytochrome P450 reductase (Fragment)	Cflo_N_g197t1	2.57	0.93	1.75	1.46	1.38
Lyso i-typ	Lysozyme	Cflo_N_g1023t1	1.18	2.21	1.70	2.35	2.16
ImdK1	Receptor-interacting serine/threonine-protein kinase 1	Cflo_N_g7081t1	2.08	0.93	1.51	1.13	1.03
Dorsal	Embryonic polarity protein dorsal	Cflo_N_g3305t1	1.80	0.97	1.38	1.15	1.10
CathD	pfam00026. Asp. Eukaryotic aspartyl protease	Cflo_N_g9172t1	1.52	1.18	1.35	0.05	†
LAP61281	Lysosomal aspartic protease (Fragment)	Cflo_N_g14504t1	1.31	1.23	1.27	3.01	2.91
zcp	Zinc carboxypeptidase A 1	Cflo_N_g8803t1	0.61	1.90	1.25	0.29	0.28
MAPKKK7	Mitogen-activated protein kinase kinase kinase 7	Cflo_N_g9451t1	1.48	0.99	1.23	1.36	1.20
TEP1	CD109 antigen	Cflo_N_g4492t1	1.30	1.10	1.20	1.38	1.33
cP45018a1	cytochrome P450 18a1	Cflo_N_g5542t1	0.78	1.58	1.18	0.28	0.28
hpchit	Putative chitin binding peritrophin-a domain protein	Cflo_N_g2215t2	0.13	2.00	1.07	0.004	0.004
TEP3	CD109 antigen	Cflo_N_g9745t1	0.98	1.11	1.04	1.20	1.15
sushi	Sushi, von Willebrand factor type A, EGF and pentraxin domain-containing protein 1	Cflo_N_g10836t1	0.92	0.90	0.91	0.48	0.46
chymo	Chymotrypsin-1	Cflo_N_g907t1	0.20	1.28	0.74	0.01	0.01
lip	Lipase member H-A (Fragment)	Cflo_N_g14858t1	0.49	0.98	0.73	0.43	0.44
chito	Acidic mammalian chitinase	Cflo_N_g2277t1	0.41	1.01	0.71	0.13	0.13
hex	Hexamerin	Cflo_N_g6501t1	0.25	1.05	0.65	1.00	0.19
hppaci	Serine protease inhibitor 3-like	Cflo_N_g1319t1	0.25	0.91	0.58	0.01	0.01

The qRT-PCR data confirmed the up- or downregulation of several genes with putative immune function (Table 10, Figure 14, Figure 15). The two genes with highest expression after immune

challenge comprise the *hypothetical protein hp67112* (Cflo_N_g6748t1; p-value p(larvae, L) = 0.0005; p(worker, W) = 0.0815) and the antimicrobial peptide gene *hymenoptaecin* (Cflo_N_g14777t1; p(L) = 0.0006; p(W) = 0.0027) with each being more than 200-fold induced upon immune challenge. Two genes, the *parathyroid hormone/ parathyroid hormone-related peptide receptor PHR* (Cflo_N_g6985t1; p(L) < 0.0001; p(W) = 0.0097), which is involved in immune signalling, and the *esterase FE4/ carboxylic-ester hydrolase* (Cflo_N_g8597t1; p(L) < 0.0001; p(W) = 0.0504, not significant), were about 21-fold more strongly expressed after immune challenge. Carboxylesterases, although mostly investigated with respect to insecticide resistance, have been shown to be involved in immune responses in *Bombyx mori*, *Galleria mellonella* and *Manduca sexta* (Serebrov et al., 2001; Shiotsuki and Kato, 1999; Zhu et al., 2003). It was proposed that carboxylesterases hydrolyse and thereby degrade toxic substances released by microbes during an infection. Furthermore, the expression of the *C. floridanus* genes encoding another AMP, defensin-1 (JN989495.1; p(L) = 0.0003; p(W) = 0.0092), the Tyrosine 3 monooxygenase (TyrOH) (Cflo_N_g5222t2; p(L) = 0.002; p(W) = 0.0001) and the alpha-2-Macroglobulin-like protein TEP2 (Cflo_N_g7345t1; p(L) = 0.0005; p(W) = 0.1354, not significant) was highly upregulated upon infection. In *Drosophila*, TyrOH expression is induced via JNK signalling (Silverman et al., 2003) and probably involved in immune-related melanisation and wound-healing, as it was proposed for *Drosophila* and *Manduca* (De Gregorio et al., 2001; Gorman et al., 2007). As antimicrobial peptides, thioester-containing proteins belong to the immune effectors expressed upon immune challenge. Interestingly, only the expression of the *C. floridanus* TEP2, which contains the functional motif CGEQ in the thioester site, was upregulated upon infection, whereas TEP1 and TEP3 showed no immune-induced regulation (Table 10).

The qRT-PCR analysis generally confirmed the results concerning immune gene expression obtained via Illumina sequencing. The expression profiles of these genes show a similar regulation as the Illumina results (up- or downregulation) in at least one of the developmental stages (larvae or adult workers). Factorial ANOVAs were applied to analyse the influence of the two effects "treatment" (immune challenge or untreated control) and "developmental stage" on the expression of the chosen immune-related genes (Appendix III, Table 23). The statistical analyses revealed that the factor "treatment" had a significant influence (p-value < 0.05) on the expression of most of the analysed genes, with the exception of the genes *cytochrome P450 18a1* (p(treatment, t) = 0.967), *zinc carboxypeptidase A 1 zcp* (p(t) = 0.5355), *lipase member H-A lip* (p(t) = 0.0896), *sushi von Willebrand factor type A* (p(t) = 0.7002), *lysosomal aspartic protease LAP61281* (p(t) = 0.5584), *vitellogenin-3 vitel* (p(t) = 0.5054), *receptor-interacting serine/threonine-protein kinase 1 ImdK1* (p(t) = 0.1502), *embryonic polarity protein dorsal* (p(t) = 0.8268), *eukaryotic aspartyl protease cathepsin D* (p(t) = 0.7508), *mitogen-activated protein kinase kinase kinase 7 MAPKKK7* (p(t) = 0.4383), *NADPH-cytochrome P450 reductase NOS1* (p(t) = 0.6798), *CD109 antigen TEP1* (p(t) = 0.3771), and *TEP3*

(p(t) = 0.9828). The expression levels of these genes were not significantly depending on the immune challenge.
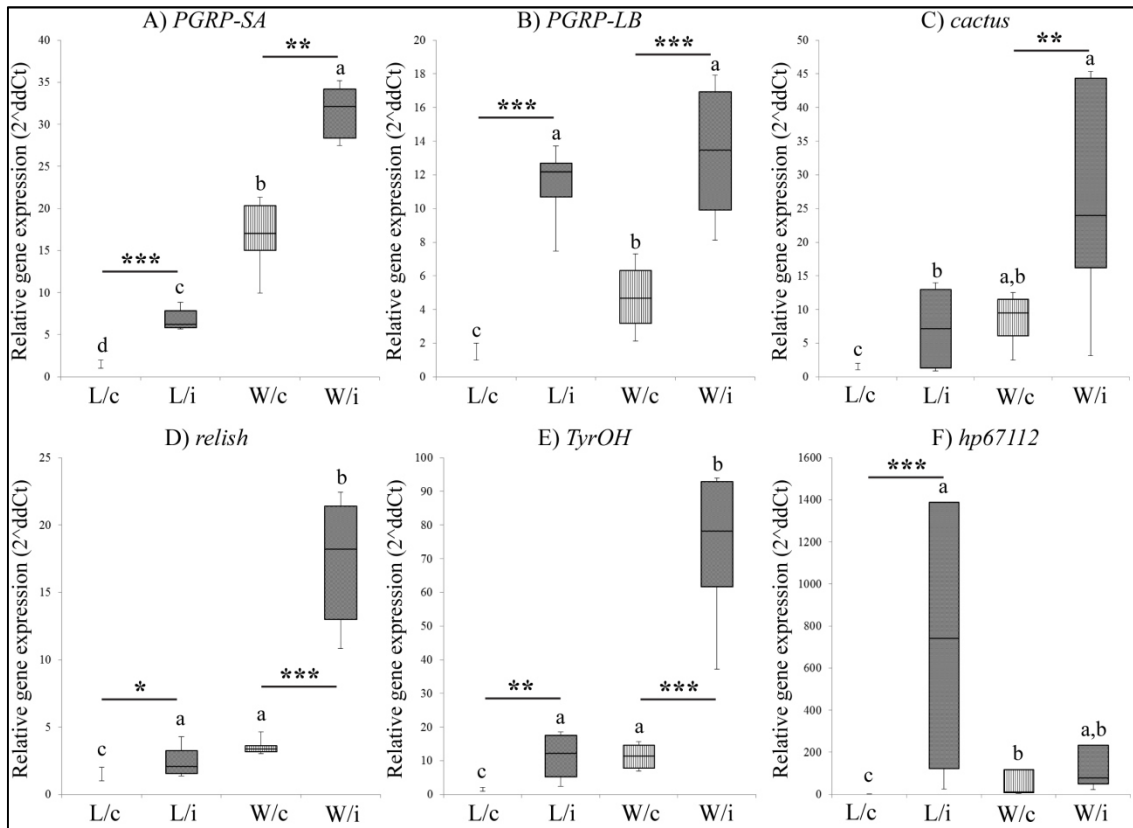
Figure 14 The expression of immune-related genes in *C. floridanus* larvae and adult workers with respect to bacterial immune challenge. *C. floridanus* larvae and workers were injected with a mixture of heat-killed *E. coli* and *M. luteus* and the expression of the immune-related genes **(A)** *peptidoglycan-recognition protein PGRP-SA*, **(B)** *PGRP-LB*, **(C)** *NF-kappa-B inhibitor cactus*, **(D)** *relish*, **(E)** *tyrosine hydroxylase* (*TyrOH*) and **(F)** *hypothetical protein hp67112* was analysed via qRT-PCR. Average gene expression levels relative to the housekeeping gene *rpL32* (dCt = Ct(target) - Ct(*rpL32*)) were determined from injected and untreated ants at 12hpi and tested for significant differences (* dependent t-test, p < 0.05; **, p < 0.01; ***, p < 0.001). Additionally, normalised changes in gene expression in the ants were tested for the significant dependency on the two factors "treatment" (immune challenge (i)/ control (c)) and "developmental stage" (larvae (L)/ workers (W)) using factorial ANOVA followed by a Tukey´s HSD post hoc test. Therefore, boxes marked with different letters (a-d) are significantly different (p < 0.05). Further results of the statistical analyses are given in Table 23 and Table 24. The Box-plots show normalised changes in gene expression relative to the expression data set with the lowest expression level (2^{ddCt}-values).

However, the separate analyses of gene expression in larvae and adults also revealed that differences in immune gene expression significantly depended on the factor "developmental stage". In fact, the expression of only three of the analysed genes was not significantly dependent on the developmental stage of the ants while being significantly dependent on the effect "treatment": *hypothetical protein hp67112* (p(t) < 0.0001; p(dv) = 0.0764), *protein yellow* (p(t) < 0.0001; p(dv) = 0.1404), and *phenoloxidase subunit A3 Posub* (p(t) = 0.0005; p(dv) = 0.1917). The expression of several genes was only downregulated in larvae, while no significant regulation was observed in workers, e.g. *chymotrypsin-1* (p(t) = 0.0092; p(dv) < 0.0001), *hexamerin* (p(t) = 0.0074; p(dv) < 0.0001) and the *putative chitin binding peritrophin-a domain protein hpchit* (p(t) = 0.008; p(dv) < 0.0001). On the other hand, although being upregulated in both, larvae and workers (Table 10), the expression of many

genes analysed appeared to be much higher in immune-challenged larvae when compared to immune-challenged workers, e.g. *hp67112* (p(t) < 0.0001; p(dv) = 0.0764, not significant), *hymenoptaecin* (p(t) < .0001; p(dv) < 0.0001), *NF-kappa-B inhibitor cactus* (p(t) < 0.0001; p(dv) = 0.0002), *TEP2* (p(t) < 0.0001; p(dv) < 0.0001), *PHR* (p(t) < 0.0001; p(dv) < 0.0001); or *PGRP-LB* (p(t) < 0.0001; p(dv) < 0.0001).
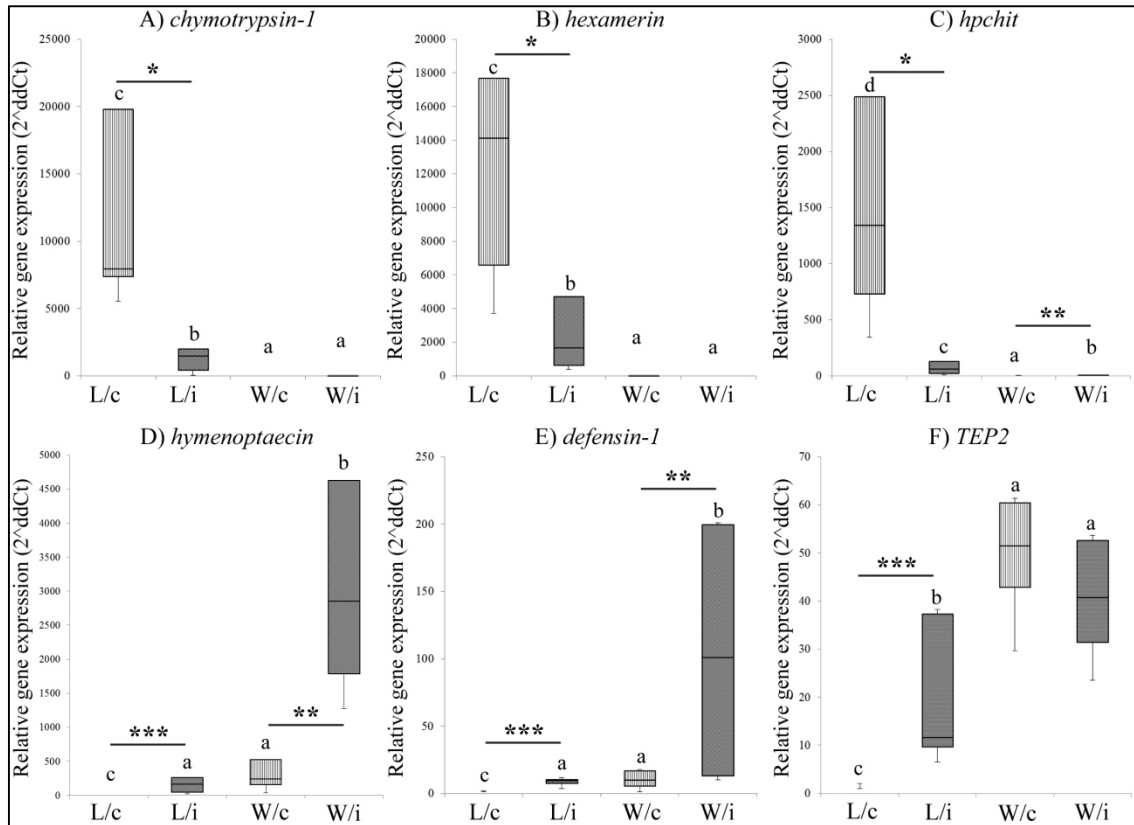

Figure 15 The expression of immune-related genes in *C. floridanus* larvae and adult workers with respect to bacterial immune challenge. *C. floridanus* larvae and workers were injected with a mixture of heat-killed *E. coli* and *M. luteus* and the expression of the genes **(A)** *chymotrypsin*, **(B)** *hexamerin*, **(C)** *putative chitin binding peritrophin-a domain protein hpchit*, **(D)** *hymenoptaecin*, **(E)** *defensin-1* and **(F)** *alpha-2-macroglobulin-like protein TEP2* was analysed via qRT-PCR. Average gene expression levels relative to the housekeeping gene *rpL32* (dCt = Ct(target) - Ct(*rpL32*)) were determined from immune-challenged and untreated ants at 12hpi and tested for significant differences (* dependent t-test, p < 0.05; **, p < 0.01; ***, p < 0.001). Additionally, normalised changes in gene expression in the ants were tested for the significant dependency on the two factors "treatment" (immune challenge (i)/ control (c)) and "developmental stage" (larvae (L)/ workers (W)) using factorial ANOVA followed by a Tukey´s HSD post hoc test. Therefore, boxes marked with different letters (a-d) are significantly different (p < 0.05). Further results of the statistical analyses are given in Table 23 and Table 24. The Box-plots show normalised changes in gene expression relative to the expression data set with the lowest expression level (2^{ddCt}-values).

Overall, the factorial ANOVAs followed by Tukey´s HSD post hoc tests (Appendix III, Table 23 and Table 24) revealed the dependency of expression levels of several immune-related genes on both effects "developmental stage" and "treatment" (Figure 14 and Figure 15). These results are in agreement with recent studies reporting differences in the expression of immune genes in various developmental stages of other holometabolous insects (Colgan et al., 2011; Fellous and Lazzaro, 2011; Randolt et al., 2008).

4.2 The haemolymph proteome and peptidome of *C. floridanus*

### 4.2.1 Haemolymph proteins in *C. floridanus*

As mentioned in the introduction, one function of the insect haemolymph is to serve as the transport medium for haemocytes, molecules and metabolites. Therefore, the haemolymph plays a role in communication, immune response, development and monitoring physiological conditions within the insects. Here, a general overview of all haemolymph proteins identified in *C. floridanus* will be given. The haemolymph of untreated and immune challenged larvae and adult workers was isolated and applied for mass spectrometry. Overall, 1,291 proteins of the new *C. floridanus* genome annotation (chapter 4.1.1) were found in the samples (about 7% of all annotated proteins). Adult workers and larvae of *C. floridanus* share 73.4% of all identified haemolymph proteins and 281 proteins were exclusively detected in the larval haemolymph samples (21.8%). Interestingly, merely 62 proteins (4.8%) were specific for the adult haemolymph samples (Figure 16).
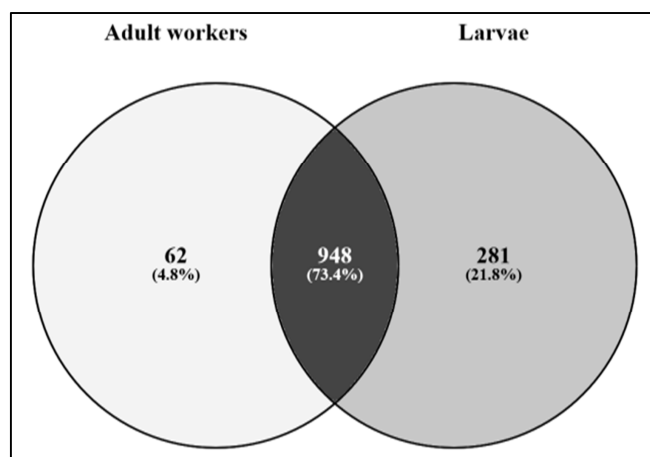
Figure 16 Venn diagram presenting razor unique proteins in the haemolymph of *C. floridanus*. The haemolymph samples of *C. floridanus* larvae and adult workers share 73.4% of all identified haemolymph proteins. 4.8% of the haemolymph proteins were detected only in adult haemolymph, whereas 21.8% of all haemolymph proteins were found only in larval haemolymph samples.

The gene ontology analyses are used to define general functions of proteins within a sample and were applied to obtain first insights into general characteristics of the haemolymph proteins regarding "biological process", "molecular function" and "cellular compartment". The full lists of haemolymph proteins and their GO terms are given in Table S1 and S2 in the digital Appendix. 721 identified haemolymph proteins were assigned to GO terms within the category "biological process", and several proteins were annotated to more than one GO term (performed by Shishir K. Gupta). The distribution of the observed GO terms is presented in Figure 17. More than one third of the assigned proteins are involved in metabolic (23.16%), protein catabolic (4.72%) and biosynthetic (7.77%) processes as well as in glycolysis (1.25%) and proteolysis (4.72%). According to the haemolymph functioning as a transport medium for cells and molecules, several proteins involved in protein transport (4.72%), lipid transport (0.83%) and iron ion transport (0.69%) were identified. Also, several proteins were assigned

to the GO terms immune response, defence response, and peptidoglycan catabolic process (2.08%), or to the GO term response to stress (0.97%).
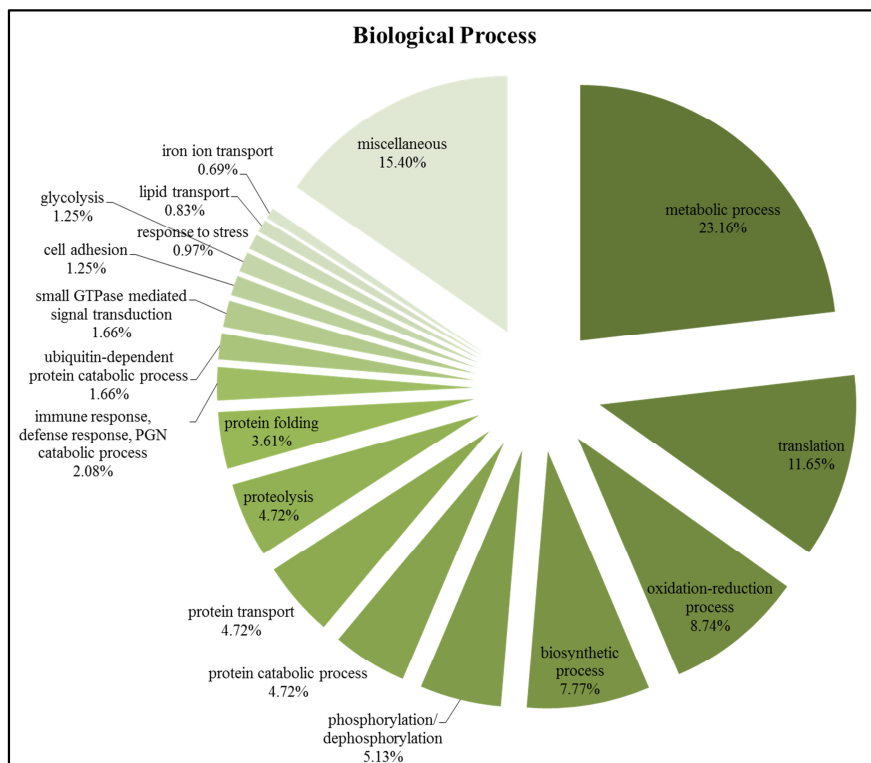

Figure 17 GO annotation for haemolymph proteins in *C. floridanus*. 721 proteins out of 1,291 identified haemolymph proteins in *C. floridanus* were assigned to GO terms for the category "biological process".

Furthermore, 450 proteins found in the haemolymph were assigned to different GO terms within the category "cellular component" (Figure 18). Almost one quarter of the proteins were categorised as proteins located within the cytoplasm or cytosol (24.44%), whereas 11.33% of the proteins belong to the extracellular region, 8.67% to ribosomes, and 7.33% to the nucleus. Also, 6.67% of the 450 proteins were categorised as membrane proteins, and 6.22% of the proteins are associated with microtubules. Interestingly several haemolymph proteins belong to organelles such as mitochondria (2.67%), the Golgi (2.22%), the endoplasmic reticulum (1.56%), synaptic vesicles (1.78%) and lysosomes (0.89%) which might indicate contamination of the samples with intracellular proteins probably occurring during injection of the bacteria and the haemolymph collection. Both procedures involve the piercing of the cuticula and skin of the insects which is likely to lead to cell damage and the leakage of intracellular proteins.

Finally, 845 proteins of the *C. floridanus* haemolymph were assigned according to their "molecular function" (Figure 19). Most of the proteins are involved in either catalytic functions or in binding processes, for example the binding of metal ions (15.15%), ATP (9.94%) and GTP (4.38%) or lipids (0.95%); proteins with hydrolase activity (11.12%), transferase activity (9.23%), oxidoreductase activity (5.33%), ligase activity (3.55%), peptidase and peptidase inhibitor activity (2.25%) and several others.
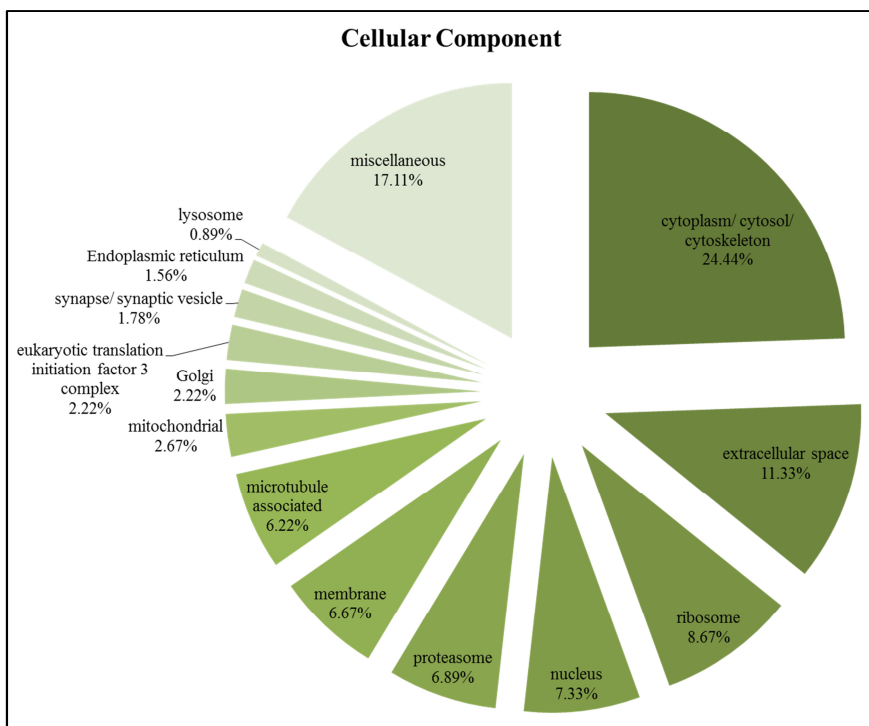
Figure 18 GO annotation for haemolymph proteins in *C. floridanus*. 450 proteins out of 1,291 identified haemolymph proteins in *C. floridanus* were assigned to GO terms for the category "cellular component".
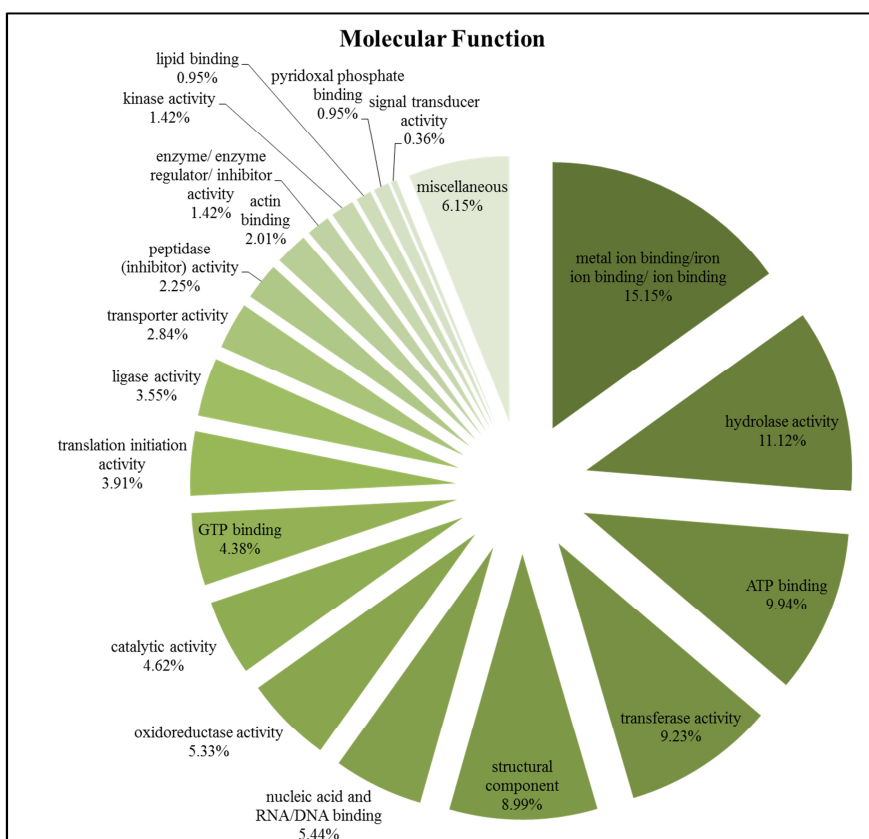
Figure 19 GO annotation for haemolymph proteins in *C. floridanus*. 845 proteins out of 1,291 identified haemolymph proteins in *C. floridanus* were assigned to GO terms for the category "molecular function".

Amongst the most abundant proteins in the haemolymph of *C. floridanus* larvae and adults were several apolipoproteins/ apolipophorins (Cflo_N_g1629.t1; Cflo_N_g13973.t1; Cflo_N_g12361.t1; Cflo_N_g13783.t1; Cflo_N_g144.t1). It was shown that insect lipoproteins, also called lipophorins, comprise two glycosylated apolipoproteins, which are core components of the lipophorin complex. A third and exchangeable component, Apolipophorin-III, is highly abundant in insect haemolymph and plays a key role in lipid transport and innate immunity (Zdybicka-Barabas and Cytryńska, 2013). Apolipophorin-III mediates pattern recognition which results in the detoxification of microbial cell wall components and in the initiation of humoral and cellular immune response mechanisms (Gupta et al., 2010; Niere et al., 2001; Whitten et al., 2004; Zdybicka-Barabas and Cytryńska, 2013). Interestingly, a homolog to other insects' Apolipophorin-III proteins was detected in the *C. floridanus* haemolymph (Cflo_N_g12361.t1). The protein annotated as Apolipophorin-III-like protein has a molecular weight of about 21 kDa and contains no cysteine residues, which is characteristic for Apolipophorin-III. Also, the protein comprises a signal peptide but shows only a low degree of amino acid sequence similarity to Apolipophorin-III proteins of other insect species (Figure 20). However, low similarities based on amino acid sequences are typical for Apolipophorin-III proteins of evolutionary divergent insect species. Rather, it was shown that the three-dimensional structures, which are connected to the Apolipophorin-III proteins' physiological function, are highly similar (Zdybicka-Barabas and Cytryńska, 2013).
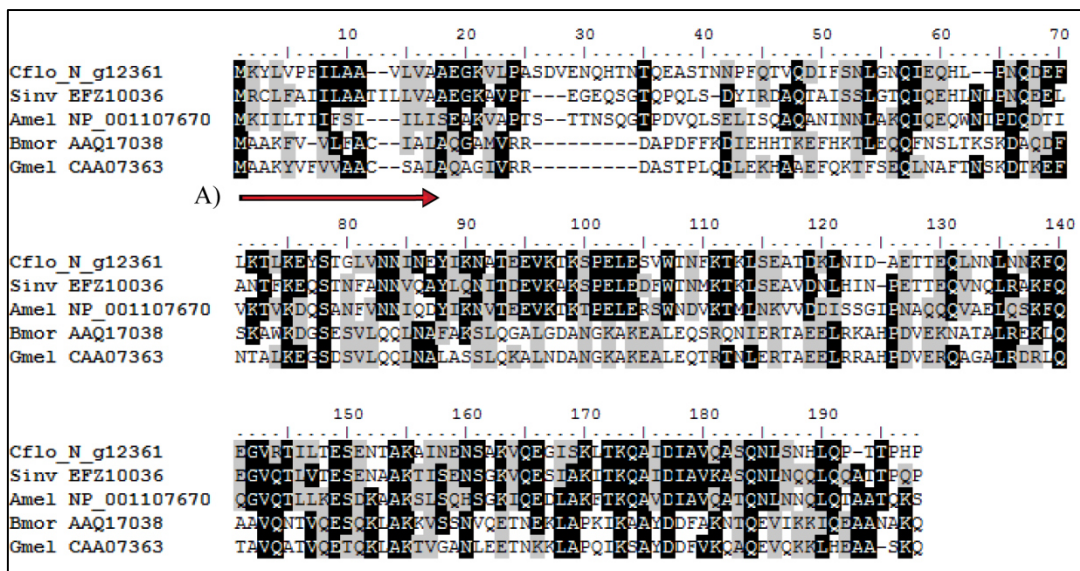
Figure 20 Alignment of the *C. floridanus* Apolipophorin-III-like protein (Cflo_N_g12361) with Apolipophorin-III proteins from *Solenopsis invicta* (Sinv; accession number EFZ10036), *Apis mellifera* (Amel; NP_001107670), *Bombyx mori* (Bmor; AAQ17038) and *Galleria mellonella* (Gmel; CAA07363). Identical amino acid residues are highlighted with a black background and similar ones with a grey background. A) The signal peptide region (about 17 to 19 amino acids in the Apolipophorin-III proteins) is marked with the red arrow.

Also, other lipoproteins, the Vitellogenin proteins, were highly abundant in the haemolymph of *C. floridanus*. In egg-laying queens of *A. mellifera* Vitellogenin proteins reach concentrations of up to 70% of the haemolymph proteins (Barchuk et al., 2002; Salmela et al., 2015). It was shown that

Vitellogenin is a yolk precursor which is synthesised in the fat body and taken up by the nurse cells and eggs (Finn, 2007). Additionally, Vitellogenin proteins serve as pattern recognition receptors binding bacterial cell wall structures or fungal glucans (Tong et al., 2010; Zhang et al., 2011). Three Vitellogenin proteins were identified in the haemolymph of *C. floridanus*. While Vitellogenin-6 (Cflo_N_g12048.t1) was highly abundant in both larvae and adults and about 64 to 71 unique peptides were counted, Vitellogenin-3 (Cflo_N_g8262.t1) showed higher abundance only in worker haemolymph (116/108 unique peptides), and Vitellogenin (Cflo_N_g6130.t1) was only detected in the larval haemolymph (135/139). In addition, the Lectin 4 C-type lectin (Cflo_N_g11262.t1) was found in the haemolymph of larvae and adult animals. C-type lectins are calcium-dependent carbohydrate-binding proteins which function in pathogen recognition, cellular interactions and innate immunity in mammals (Weis et al., 1998; Yu et al., 2002). However, C-type lectins are also present in insects and lepidopteran lectins have been reported to participate in innate immune responses such as phagocytosis and activation of prophenoloxidases (Jomori and Natori, 1992; Yu and Kanost, 2000; Yu et al., 2002). In *Manduca sexta* and *Bombyx mori* the so-called immulectins (IML) contain two tandem C-type lectin carbohydrate recognition domains (Koizumi et al., 1999; Yu and Kanost, 2000). The Lectin 4 C-type lectin identified in the haemolymph of *C. floridanus* contains one C-type lectin (CTL)/ C-type lectin-like (CTLD) domain including the ligand binding site and therefore could be involved in recognition of MAMPs in *C. floridanus*. One of the most abundant proteins in all haemolymph samples was the Hemocytin (Cflo_N_g11181.t1). The protein is a homolog to the *B. mori* Hemocytin which is an ortholog of the mammalian von Willebrand factor. It was shown that hemocytins from the silkworm *B. mori* and *Drosophila* are initiators of nodulation (Arai et al., 2013; Lavine and Strand, 2002).

Further, proteins involved in metabolism and storage were identified in high concentrations in the haemolymph of *C. floridanus* larvae and workers, for example the alpha-N-Acetylgalactosaminidase (Cflo_N_g15021.t1), the Aminoacylase-1-like protein (Cflo_N_g1641.t1), the Cytosol aminopeptidase (Cflo_N_g8453.t1), the Glycogen debranching enzyme (Cflo_N_g9966.t1), or the Glyceraldehyde-3-phosphate dehydrogenase (Cflo_N_g7809.t1), and proteins involved in glycolysis, such as Enolase (Cflo_N_g10715.t1), the Phosphoglycerate mutase 1 (Cflo_N_g13024.t1), and the Fructose-bisphosphate aldolase (Cflo_N_g1444.t1). In addition, five Hexamerin proteins were identified within the *C. floridanus* haemolymph. Four Hexamerin proteins (Cflo_N_g10613.t1; Cflo_N_g6501.t1; Cflo_N_g10612.t1 and Cflo_N_g6500.t1) were highly abundant in the haemolymph of larvae and adults, whereas the fifth Hexamerin (Cflo_N_g11205.t1) reached higher concentrations in larval haemolymph (21/22 unique peptides in larvae when compared to adult haemolymph with 9/10 unique peptides). Hexamerin proteins are hemocyanin-derived proteins participating in the storage of amino acids in various insects (Capurro et al., 2000; Martinez et al., 2000; Martins et al., 2010; Wheeler and Martinez, 1995).

Results

Six proteins orthologous to glutathione S-transferases (GSTs) were found in the *C. floridanus* haemolymph in high concentrations: two Glutathione S-transferases (Cflo_N_g10071.t1; Cflo_N_g10070.t1), a putative Glutathione-S-transferase Theta (Cflo_N_g7328.t1), a Glutathione S-transferase 1, isoform D-like protein (Cflo_N_g2628.t1), a Glutathione S-transferase 1-like protein (Cflo_N_g6258.t1) and the Glutathione S-transferase Omega-1 protein (Cflo_N_g4685.t2). It was shown that GSTs are widely distributed in insects participating in the enzymatical degradation of various insecticides and therefore mediate insecticide tolerance or resistance (Fang, 2012; Kostaropoulos et al., 2001a, b). Several other proteins involved in response to stress were identified in the haemolymph, e.g. the Heat shock 70 kDa protein Cognate 4-like isoform 2 (Cflo_N_g3612.t1), Heat shock protein beta-1 (Cflo_N_g1474.t2), Heat shock protein 90 (Cflo_N_g7863.t1) or the Glutathione peroxidase (Cflo_N_g2140.t1). Moreover, several proteins involved in iron transport, storage and regulation were detected: two Transferrins (Cflo_N_g7714.t1; Cflo_N_g2636.t1), the Ferritin 2-like protein (Cflo_N_g1204.t1), a Ferritin subunit (Cflo_N_g1203.t1) and a protein annotated as Soma Ferritin (Cflo_N_g10406.t1). It was shown that insect ferritins play a role in iron homeostasis (Tang and Zhou, 2013) and in the protection against oxidative stress (Pham and Winzerling, 2010).

Several proteins with previously described immune function were found in the haemolymph of the ants. These proteins include the PGN recognition proteins PGRP-SA (Cflo_N_g8526.t1), PGRP-SC2 (Cflo_N_g102.t2) and PGRP-LB (Cflo_N_g103.t1) and the alpha-2-Macroglobulin-like protein TEP2 (Cflo_N_g7345.t1). Additionally, two PRRs probably involved in recognition of Gram-positive bacteria, the beta-1,3-glucan-binding proteins GNBP and GNBP1_3 (Cflo_N_g15215.t1 and Cflo_N_g5742.t1), were identified. Furthermore, two Niemann-Pick C2 (NPC2)-like proteins (Cflo_N_g1858.t1 and Cflo_N_g7354.t1), were present in the haemolymph of adult workers and larvae. In *D. melanogaster*, NPC2 proteins bind to bacterial cell wall structures, and it was suggested that NPC2 proteins may be involved in the IMD signalling pathway (Shi et al., 2012). Also, a Phenoloxidase subunit A3 (Cflo_N_g1918.t1) and the serine proteinase Stubble (Cflo_N_g14403.t1) were found in the haemolymph samples. The latter is a homolog of the Serine protease homolog 42 from *A. mellifera* (GenBank Acc. No.: XP_623150) which was shown to be involved in prophenoloxidase cleavage and activation upon septic injury (Zou et al., 2006). Therefore, it is likely that the *C. floridanus* serine protease Stubble participates in the initiation of melanisation processes. Several components of the Toll signalling pathway were identified within the ants' haemolymph, such as the serine protease Snake 2 (Cflo_N_g7945.t1) and two proteins annotated as protein Spätzle (Cflo_N_g12735.t1, Cflo_N_g9642.t1) with the latter found in the adult haemolymph only. Also two serine proteases Persephone (Cflo_N_g8446.t1 and Cflo_N_g8442.t1) were detected.

Furthermore, various putative effectors of the ant's immune system were detected in the haemolymph samples. A i-type Lysozyme (Cflo_N_g1023.t1) and two c-type Lysozyme c-1 proteins

(Cflo_N_g4036.t1 and Cflo_N_g5519.t1) were identified, with the latter being present only in the haemolymph of adult animals. Also, the two antimicrobial peptides defensin-1 (JN989495.1) and hymenoptaecin (Cflo_N_g14777.t1) were found in the *C. floridanus* haemolymph. Interestingly, another putative effector protein in the defence response to bacteria, the waprin-Phi1 (Cflo_N_g4956.t1), was discovered in the haemolymph. In 2007, the recombinant protein Omwaprin (the waprin from the inland taipan *Oxyuranus microlepidotus)* was expressed in *E. coli* and purified. The protein showed antimicrobial activity against the two Gram-positive bacteria *Bacillus megaterium* and *Staphylococcus warneri* (Nair et al., 2007).

4.2.2 Regulation of *C. floridanus* haemolymph proteins upon infection

In order to identify haemolymph proteins with putative immune functions in *C. floridanus* larvae and adult workers were injected with heat-killed bacteria and haemolymph samples were taken at three time points after infection. Therefore, a qualitative overview of haemolymph proteins present in the haemolymph of untreated and immune-challenged ants was obtained. Additionally, a semi-quantitative analysis regarding the enrichment of proteins in the haemolymph after immune challenge was performed. Proteins were considered to be enriched with a log2(x-fold change) of $\geq$ 1 or depleted with a log2(x-fold change) of $\leq$ -1.

Overall, 391 proteins were enriched or depleted in the haemolymph of immune-challenged workers when compared to control samples. Out of these proteins, 75 were significantly enriched or depleted. In the haemolymph of immune-challenged larvae, 347 proteins were found to be either enriched or depleted, with 126 significantly regulated proteins (Figure 21). Only 12 proteins were significantly regulated (enriched or depleted) in both, larval and adult haemolymph upon immune challenge (Table 11), whereas the significant and non-significant enrichment or depletion of several proteins involved in the immune response appeared to be developmental stage-dependent (Table 12; Table 13 and Table 14; digital Appendix Table S1).
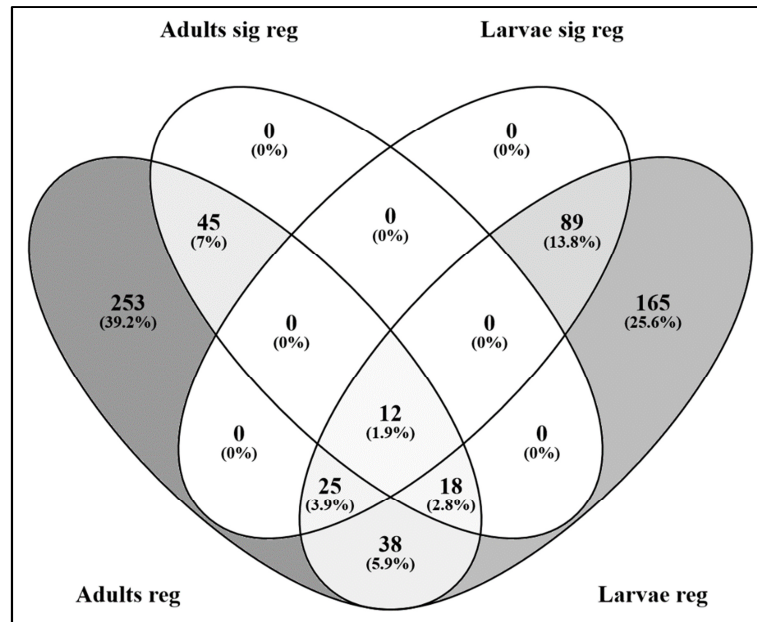
Figure 21 The distribution of 645 enriched and depleted proteins in the haemolymph of immune-challenged *C. floridanus* presented in a Venn diagram. The haemolymph samples of workers (Adults reg or Adults sig reg) and larvae (Larvae reg or Larvae sig reg) share 93 enriched or depleted proteins with only 12 proteins being either significantly enriched or depleted (sig reg) upon infection in both developmental stages.

Table 11 List of proteins significantly enriched or depleted in both larval and adult haemolymph upon immune challenge in *C. floridanus*. Enrichment and depletion are given as log2(x-fold changes), whereas significance levels are defined as significant (1) and highly significant (2).

Acc. No. | protein name	log2(x-fold change) adults	Log2(x-fold change) larvae	Significance level (adults injected)	Significance level (larvae injected)
Cflo_N_g665.t1|Hypothetical protein	5.01	4.92	1	2
Cflo_N_g8574.t1|Tropomyosin	4.91	3.08	1	1
Cflo_N_g938.t1|Glycine cleavage system H protein, mitochondrial	4.70	4.54	1	2
Cflo_N_g7789.t1|Hypothetical protein	4.20	3.94	1	1
Cflo_N_g9457.t1|Enoyl-CoA hydratase domain-containing protein 3, mitochondrial	4.11	6.66	1	2
Cflo_N_g2134.t1|Troponin C, isoform 3	3.83	2.31	1	1
Cflo_N_g1858.t1|Protein NPC2-like protein	3.34	-2.63	1	1
Cflo_N_g14697.t1|Hypothetical protein	3.22	4.20	1	1
Cflo_N_g10982.t1|titin isoform X14	3.09	3.42	1	2
Cflo_N_g5638.t1|Myosin light chain alkali	2.94	2.52	1	1
Cflo_N_g6298.t1|Zinc metalloprotease zmpB	-2.57	-3.30	1	1
Cflo_N_g5481.t1|Beta-galactosidase (Fragment)	-2.68	-2.79	1	1

Whereas two putative pattern recognition receptors (PRRs), the Hemolymph lipopolysaccharide-binding protein (Cflo_N_g10889.t1; log2(x-fold change) adults (a) = 1.35; log2(x-fold change) larvae (l) = 1.73) and PGRP-LB (Cflo_N_g103.t1; a = 2.28; l = 1.74) were enriched in larval and worker haemolymph after immune challenge, the PGRP-SA (Cflo_N_g8526.t1; l = 1.56) was only enriched in larval haemolymph. A third PGRP, the PGRP-SC2 (Cflo_N_g102.t2; a = 1.17), and the GNBP

(Cflo_N_g5742.t1; a = 1.82) were enriched in the haemolymph of immune-challenged workers only. The alpha-2-Macroglobulin-like protein TEP2 (Cflo_N_g7345.t1; l = 2.31), which supposedly binds to bacterial surfaces, was enriched in the haemolymph of immune-challenged larvae. Additionally, the Lectin 4 C-type lectin (Cflo_N_g11262.t1; a = 2.95; l = 1.16) was enriched in the haemolymph of adults and larvae. These results also confirm data from the differential expression analysis of immune genes in *C. floridanus* (chapter 4.1.4), as the expression of the genes *PGRP-LB* (Cflo_N_g103.t1) and the *GNBP* (Cflo_N_g5742.t1) was highly upregulated after an immune challenge (Gupta et al., 2015).

Interestingly, a NPC2-like protein (Cflo_N_g1858.t1) was significantly and highly enriched in adult haemolymph (log2(x-fold change) = 3.34) but significantly depleted in larval haemolymph (log2(x-fold change) = -2.63) upon immune challenge (Table 11). In contrast, another NPC2-like protein (Cflo_N_g7354.t1) was significantly enriched upon infection only in larval haemolymph (log2(x-fold change) = 2.11) (Table 13). In adult haemolymph the latter NPC2-like protein was non-significantly but about 2-fold enriched. As suggested for the *Drosophila* NPC2 proteins (Shi et al., 2012), the *C. floridanus* homologs might be involved in recognition of bacterial cell wall structures and may therefore be enriched in the haemolymph of immune-challenged animals.

Table 12 List of proteins enriched or depleted in haemolymph samples from immune-challenged *C. floridanus* workers when compared to control haemolymph. Enrichment and depletion are given as log2(x-fold changes), whereas significance levels are defined as significant (1) and highly significant (2). The respective log2(x-fold changes)-values of the proteins in larval haemolymph are given in the last column. In larvae, these proteins are non-significantly enriched or depleted, or not regulated (n.r.).

acc # plus annotation	log2(x-fold change) adults injected	Significance level (adults injected)	Log2(x-fold change) larvae injected
Cflo_N_g9025.t1|3-hydroxyacyl-CoA dehydrogenase type-2	6.95	2	1.89
Cflo_N_g11679.t1|Succinyl-CoA ligase [ADP-forming] subunit beta, mitochondrial	6.61	2	n.r.
Cflo_N_g5429.t1|Succinate dehydrogenase [ubiquinone] iron-sulfur subunit, mitochondrial	6.54	2	n.r.
Cflo_N_g13480.t1|Circadian clock-controlled protein	6.18	2	2.78
Cflo_N_g11204.t1|Troponin C	5.95	2	n.r.
Cflo_N_g8048.t1|Electron transfer flavoprotein subunit beta	5.52	2	n.r.
Cflo_N_g961.t1|Pyruvate carboxylase, mitochondrial	5.26	1	n.r.
Cflo_N_g12886.t1|protein CREG1-like	5.19	2	n.r.
Cflo_N_g13613.t1|Protein bicaudal D	5.15	2	n.r.
Cflo_N_g13165.t1|Putative GTP cyclohydrolase 1 type 2	5.08	1	n.r.
Cflo_N_g4390.t1|probable salivary secreted peptide-like	4.96	1	1.52
Cflo_N_g13028.t1|ABC transporter F family member 4-like isoform X1	4.92	1	n.r.
Cflo_N_g7037.t1|Programmed cell death protein 6	4.83	1	n.r.
Cflo_N_g3789.t1|Tube	4.63	1	n.r.
Cflo_N_g8933.t1|PDZ and LIM domain protein Zasp-like isoform X8	4.61	1	n.r.
Cflo_N_g9198.t1|transport and Golgi organization 2 homolog	4.58	1	n.r.
Cflo_N_g11588.t1|Succinyl-CoA ligase [GDP-forming] subunit	4.56	1	1.01

alpha, mitochondrial			
Cflo_N_g3359.t1|Thioredoxin-dependent peroxide reductase, mitochondrial	4.55	1	n.r.
Cflo_N_g1862.t1|V-type proton ATPase subunit D 1	4.37	1	n.r.
Cflo_N_g7792.t1|GRP multi-domain protein	4.26	1	-2.42
Cflo_N_g1992.t1|Hypothetical protein	4.24	1	n.r.
Cflo_N_g8908.t1|Gamma-interferon-inducible lysosomal thiol reductase	4.17	1	n.r.
Cflo_N_g4202.t1|Dihydrolipoyl dehydrogenase	4.08	1	n.r.
Cflo_N_g11020.t1|Thioredoxin	4.01	1	n.r.
Cflo_N_g12068.t1|Circadian clock-controlled protein	3.82	1	n.r.
Cflo_N_g7211.t2|Aconitate hydratase, mitochondrial	3.81	1	n.r.
Cflo_N_g4055.t1|sex-regulated protein janus-A-like	3.74	1	n.r.
Cflo_N_g11783.t1|SMC_N multi-domain protein	3.66	1	2.45
Cflo_N_g13068.t1|myosin regulatory light chain 2-like	3.61	1	1.96
Cflo_N_g4041.t1|GI23607	3.53	1	n.r.
Cflo_N_g10931.t1|Glutamate dehydrogenase, mitochondrial	3.49	1	1.19
Cflo_N_g1634.t1|Dehydrogenase/reductase SDR family member 4	3.46	1	-1.10
Cflo_N_g9000.t1|Ufm1-conjugating enzyme 1	3.45	1	-1.00
Cflo_N_g687.t2|Muscle LIM protein Mlp84B	3.43	1	1.19
Cflo_N_g1095.t1|Superoxide dismutase	3.39	1	1.08
Cflo_N_g9921.t1|THAP domain-containing protein 4	3.38	1	n.r.
Cflo_N_g9141.t1|pfam00168, C2, C2 domain	3.37	1	1.99
Cflo_N_g5132.t2|Troponin T	3.35	1	n.r.
Cflo_N_g12920.t1|Eukaryotic translation initiation factor 5A, putative	3.30	1	n.r.
Cflo_N_g9607.t1|Tropomyosin-1	3.28	1	1.20
Cflo_N_g10678.t1|Muscle-specific protein 20	3.21	1	1.29
Cflo_N_g12294.t1|Cytochrome c-2	3.19	1	n.r.
Cflo_N_g12762.t2|Alanine aminotransferase 2	3.03	1	n.r.
Cflo_N_g8135.t1|Troponin C	2.97	1	n.r.
Cflo_N_g11262.t1|Lectin 4 C-type lectin	2.95	1	1.16
Cflo_N_g14387.t1|unclassified	2.90	1	n.r.
Cflo_N_g7919.t1|Globin	2.86	1	n.r.
Cflo_N_g515.t1|Amphiphysin	2.76	1	n.r.
Cflo_N_g1023.t1|i-type Lysozyme	2.64	1	n.r.
Cflo_N_g3939.t1|Electron transfer flavoprotein subunit alpha, mitochondrial	2.61	1	n.r.
Cflo_N_g596.t1|Peroxisomal membrane protein PEX16	2.60	1	n.r.
Cflo_N_g13055.t1|Fumarate hydratase, mitochondrial	2.57	1	n.r.
Cflo_N_g2300.t1|Bifunctional 3'-phosphoadenosine 5'-phosphosulfate synthetase	-2.67	1	-1.23
Cflo_N_g4956.t1|Waprin-Phi1	-2.78	1	n.r.
Cflo_N_g9231.t1|DE-cadherin	-2.95	1	n.r.
Cflo_N_g7519.t1|Catalase	-2.99	1	n.r.
Cflo_N_g14552.t1|Alcohol dehydrogenase [NADP+]	-3.00	1	n.r.
Cflo_N_g578.t1|Beta-ureidopropionase	-3.34	1	n.r.
Cflo_N_g9571.t1|Teneurin-3	-3.42	1	n.r.
Cflo_N_g10398.t1|Phosphoglycerate kinase	-3.92	1	n.r.
Cflo_N_g9828.t1|Hypothetical protein	-3.96	1	1.27
Cflo_N_g13339.t1|Adenosine kinase 2	-4.27	1	n.r.
Cflo_N_g13094.t1|Larval cuticle protein A1A	-4.78	2	n.r.

Several regulatory components of the Toll signalling pathway were enriched in the immune challenged haemolymph samples, e.g. the serine proteases Snake 2 (Cflo_N_g7945.t1; a = 1.82;

l = 1.88) and Persephone (Cflo_N_g8446.t1; a = 2.04; l = 1.48). Another serine protease Persephone (Cflo_N_g8442.t1; l = 2.03) was enriched only in larvae whereas two proteins annotated as Spätzle (Cflo_N_g9642.t1; a = 1.08 and Cflo_N_g12735.t1; a = 1.04) were enriched in the haemolymph of immune-challenged workers. Moreover, putative immune effectors, such as the AMPs defensin-1 (JN989495.1; a = 2.55; l = 3.32) and hymenoptaecin (Cflo_N_g14777.t1; a = 2.45; l = 4.70), were enriched in larval and worker haemolymph, whereas i-type Lysozyme (Cflo_N_g1023.t1; a = 2.64) was enriched in haemolymph of immune-challenged adults only. Interestingly, only the expression of genes encoding serine protease Persephone (Cflo_N_g8442.t1) and defensin-2 (Cflo_N_g8312.t1) was upregulated upon infection (Gupta et al., 2015).

Table 13 List of proteins enriched or depleted in haemolymph samples of immune-challenged larvae. Enrichment and depletion are given as log2(x-fold changes), whereas significance niveaus are defined as significant (1) and highly significant (2). The respective log2(x-fold changes)-values of the proteins in adult haemolymph are given in the last column. In adult workers these proteins are non-significantly enriched or depleted, or not regulated (n.r.).

acc # plus annotation	log2(x-fold change) larvae injected	Significance level (larvae injected)	log2(x-fold change) adults injected
Cflo_N_g10681.t1|Chymotrypsin inhibitor	6.91	2	1.35
Cflo_N_g13463.t1|Hypothetical protein	4.92	2	-0.22
Cflo_N_g3051.t1|Chitin binding protein	4.86	2	-1.72
Cflo_N_g14777.t1|Hymenoptaecin	4.70	2	2.45
Cflo_N_g12019.t1|Sec61 beta family protein	4.56	2	n.r.
Cflo_N_g1121.t1|Mago nashi	4.31	1	n.r.
Cflo_N_g4671.t1|Myosin-2 essential light chain	4.26	1	n.r.
Cflo_N_g7624.t1|Coatomer subunit delta	4.26	1	n.r.
Cflo_N_g1882.t1|Ribosomal protein L19	4.16	1	n.r.
Cflo_N_g8574.t3|Tropomyosin	4.11	1	1.99
Cflo_N_g4958.t1|[O] COG2214 DnaJ-class molecular chaperone	4.03	1	n.r.
Cflo_N_g3050.t1|Putative peritrophin	4.01	2	n.r.
Cflo_N_g78.t1|Poly [ADP-ribose] polymerase	3.79	1	n.r.
Cflo_N_g8158.t1|Acidic mammalian chitinase	3.75	1	n.r.
Cflo_N_g8193.t1|Placental protein 11	3.73	1	2.86
Cflo_N_g5984.t1|transmembrane protease serine 9	3.69	1	n.r.
Cflo_N_g6471.t1|MIP-.t3 multi-domain protein	3.59	1	1.00
Cflo_N_g13627.t1|Glycerol-3-phosphate dehydrogenase [NAD+], cytoplasmic	3.55	1	n.r.
Cflo_N_g7743.t1|TPPP family protein CG4893	3.50	1	1.89
Cflo_N_g12788.t1|U6 snRNA-associated Sm-like protein LSm1	3.43	1	n.r.
Cflo_N_g12187.t1|Silk fibroin	3.43	2	n.r.
Cflo_N_g6055.t1|Clathrin heavy chain	3.38	2	n.r.
Cflo_N_g8698.t2|spectrin beta chain isoform X2	3.38	1	-2.08
JN989495.1|Defensin-1	3.32	2	2.55
Cflo_N_g11525.t1|Vacuolar protein sorting-associated protein 28-like protein	3.24	1	n.r.
Cflo_N_g9204.t3|DNA ligase 1-like isoform X6	3.23	1	1.76
Cflo_N_g13568.t1|Histone deacetylase complex subunit SAP18	3.21	1	n.r.
Cflo_N_g13653.t2|protein SZ.t2-like	3.20	1	n.r.

Cflo_N_g5165.t1|similar to CG18431 CG18431-PA	3.12	2	-1.88
Cflo_N_g1851.t1|unclassified	3.00	1	n.r.
Cflo_N_g11629.t1|Putative odorant binding protein	2.99	2	1.31
Cflo_N_g8597.t1|Esterase FE4; AltName: Full=Carboxylic-ester hydrolase; Flags: Precursor	2.99	2	n.r.
Cflo_N_g4257.t1|Leukocyte elastase inhibitor	2.97	1	n.r.
Cflo_N_g1411.t1|Paramyosin, short form	2.94	1	n.r.
Cflo_N_g9424.t1|proteoglycan 4-like isoform X1	2.91	1	n.r.
Cflo_N_g11378.t1|Clathrin light chain	2.90	1	n.r.
Cflo_N_g12890.t2|Myosin heavy chain, muscle	2.89	1	n.r.
Cflo_N_g12341.t1|saccharopine dehydrogenase-like oxidoreductase-like, partial	2.73	1	n.r.
Cflo_N_g3658.t2|microtubule-associated protein Jupiter-like isoform 1	2.66	1	n.r.
Cflo_N_g9577.t1|GMP reductase	2.50	1	2.21
Cflo_N_g1853.t1|mucin-6-like	2.46	1	n.r.
Cflo_N_g10682.t1|chymotrypsin inhibitor-like, partial	2.44	1	n.r.
Cflo_N_g10280.t1|Pancreatic triacylglycerol lipase (Fragment)	2.42	1	2.55
Cflo_N_g10198.t1|Translation initiation factor IF-2	2.42	1	2.47
Cflo_N_g9425.t1|heterochromatin protein 1-binding protein 3-like	2.38	1	-1.11
Cflo_N_g7345.t1|Alpha-2-macroglobulin	2.31	1	n.r.
Cflo_N_g1735.t1|collagen alpha-1(IV) chain	2.29	1	n.r.
Cflo_N_g12524.t1|15-hydroxyprostaglandin dehydrogenase [NAD+]	2.26	1	n.r.
Cflo_N_g9831.t1|Obstractor C1	2.23	1	n.r.
Cflo_N_g7130.t1|Putative odorant-binding protein A10	2.19	1	n.r.
Cflo_N_g10679.t2|Mitogen-activated protein kinase kinase kinase 15	2.12	1	n.r.
Cflo_N_g7354.t1|Protein NPC2-like protein	2.11	1	1.12
Cflo_N_g2144.t2|GATA zinc finger domain-containing protein 14-like	2.09	1	n.r.
Cflo_N_g8442.t1|Serine protease persephone	2.03	1	n.r.
Cflo_N_g8989.t1|Dynein beta chain, ciliary	2.01	1	n.r.
Cflo_N_g10706.t1|FK506-binding protein 2-like	1.96	1	n.r.
Cflo_N_g12496.t2|Titin	1.92	1	n.r.
Cflo_N_g10896.t1|growth/differentiation factor 11-like	1.91	1	2.55
Cflo_N_g7887.t1|Hypothetical protein	1.90	1	n.r.
Cflo_N_g2156.t1|chymotrypsin 2 [Anopheles gambiae]	1.82	1	n.r.
Cflo_N_g103.t1|PGRP-LB	1.74	1	2.28
Cflo_N_g10729.t1|PAB-dependent poly(A)-specific ribonuclease subunit 3	1.70	1	n.r.
Cflo_N_g793.t1|Calexcitin-2	-1.70	1	n.r.
Cflo_N_g6223.t1|Placental protein 11	-1.71	1	n.r.
Cflo_N_g6564.t1|Fatty acid-binding protein	-1.77	1	n.r.
Cflo_N_g8803.t1|Zinc carboxypeptidase A 1	-1.80	1	n.r.
Cflo_N_g14527.t1|Carbonic anhydrase 2	-1.89	1	1.65
Cflo_N_g7296.t1|Chymotrypsin-1	-1.89	1	n.r.
Cflo_N_g293.t2|Glucosamine--fructose-6-phosphate aminotransferase [isomerizing] 2	-2.02	1	n.r.
Cflo_N_g11635.t1|Guanine deaminase	-2.08	1	1.28
Cflo_N_g5287.t1|Aminopeptidase N	-2.14	1	n.r.
Cflo_N_g9539.t1|Acyl-coenzyme A:6-aminopenicillanic-acid-acyltransferase 40 kDa form	-2.14	1	n.r.
Cflo_N_g2544.t1|Trypsin-1	-2.17	1	1.78
Cflo_N_g11991.t1|Ribosomal protein L27Ae	-2.25	1	n.r.
Cflo_N_g14742.t1|Guanine deaminase	-2.26	1	n.r.
Cflo_N_g559.t1|Charged multivesicular body protein 4b	-2.27	1	n.r.

Cflo_N_g3612.t1|heat shock 70 kDa protein cognate 4-like isoform 2	-2.31	1	n.r.
Cflo_N_g8960.t1|Lysosomal alpha-mannosidase	-2.36	1	-1.96
Cflo_N_g15545.t1|Maltase 2	-2.38	1	n.r.
Cflo_N_g11628.t1|unclassified	-2.44	1	n.r.
Cflo_N_g2490.t1|Alkaline phosphatase	-2.46	1	n.r.
Cflo_N_g5252.t1|Protein G12	-2.57	1	n.r.
Cflo_N_g3414.t1|Acetyl-CoA carboxylase	-2.62	1	n.r.
Cflo_N_g5290.t1|Aminopeptidase N	-2.64	1	n.r.
Cflo_N_g12700.t1|Protein Red	-2.71	1	n.r.
Cflo_N_g12121.t1|Prostatic acid phosphatase	-2.72	1	n.r.
Cflo_N_g13757.t1|pfam02958, EcKinase, Ecdysteroid kinase	-2.72	1	n.r.
Cflo_N_g1453.t1|Alpha-glucosidase	-2.75	1	n.r.
Cflo_N_g7375.t1|Importin-5	-2.87	1	n.r.
Cflo_N_g9721.t1|Peritrophin-48 (Fragment)	-2.90	2	n.r.
Cflo_N_g3601.t1|60S ribosomal protein L14	-3.03	2	n.r.
Cflo_N_g3583.t1|Protein 5NUC	-3.04	1	n.r.
Cflo_N_g5983.t1|Biotin synthesis protein bioC	-3.10	1	n.r.
Cflo_N_g4486.t1|Glucosylceramidase	-3.10	1	n.r.
Cflo_N_g4444.t1|trypsin-1-like	-3.13	2	n.r.
Cflo_N_g5236.t1|Protein G12	-3.18	2	n.r.
Cflo_N_g14017.t1|Xanthine dehydrogenase	-3.22	1	-1.31
Cflo_N_g6687.t1|Maltase 1	-3.30	1	n.r.
Cflo_N_g5611.t1|Probable glucosamine 6-phosphate N-acetyltransferase	-3.38	1	n.r.
Cflo_N_g522.t1|Ras GTPase-activating protein-binding protein 2	-3.53	1	n.r.
Cflo_N_g6124.t1|Peritrophin-1	-3.62	2	n.r.
Cflo_N_g11664.t1|Galectin (Fragment)	-3.66	1	n.r.
Cflo_N_g10976.t1|Esterase E4; AltName: Full=Carboxylic-ester hydrolase; Flags: Precursor	-3.74	2	n.r.
Cflo_N_g4217.t1|niemann-Pick C1 protein-like	-3.77	1	n.r.
Cflo_N_g13662.t1|pfam06585, JHBP, Haemolymph juvenile hormone binding protein (JHBP)	-3.80	1	n.r.
Cflo_N_g4555.t1|Peritrophic membrane protein 1 (Fragment)	-3.86	2	n.r.
Cflo_N_g8234.t1|Maltase 1	-3.86	2	n.r.
Cflo_N_g5251.t1|Protein G12	-4.45	2	n.r.
Cflo_N_g4442.t1|Trypsin-1	-4.51	2	n.r.
Cflo_N_g4317.t1|Peritrophin-1	-4.54	2	n.r.
Cflo_N_g10821.t1|Putative serine protease K12H4.7	-4.94	1	n.r.
Cflo_N_g7901.t1|Hypothetical protein	-5.14	2	n.r.
Cflo_N_g9225.t1|Ribosomal protein L15	-5.37	2	n.r.
Cflo_N_g13877.t1|Cytoplasmic dynein 2 light intermediate chain 1	-7.82	2	n.r.

The Vitellogenin-3 protein (Cflo_N_g8262.t1; $a = 1.42$; $l = -1.24$) was non-significantly depleted in larval haemolymph but enriched in adult haemolymph samples after immune challenge. Accordingly, the expression of the respective gene encoding Vitellogenin-3 was upregulated in immune-challenged *C. floridanus* workers (Gupta et al., 2015). Also, the Glutathione S-transferase 1-like protein (Cflo_N_g6258.t1; $a = 2.00$) was enriched in the haemolymph of immune-challenged adults. Hemocytin (Cflo_N_g11181.t1; $a = 1.01$), a homolog to the silkworm Hemocytin, which was suggested to play a role in immune-induced nodulation (Arai et al., 2013), was enriched in the haemolymph of *C. floridanus* workers after immune challenge. Two Ferritin proteins showed regulation upon immune challenge in larvae (Table 14). Whereas the Ferritin (Cflo_N_g1203.t1; $l = $ -

1.21) was depleted, the Soma Ferritin (Cflo_N_g10406.t1; l = 1.32) was enriched in the haemolymph. Finally, an additional putative AMP, the waprin-Phi1 (Cflo_N_g4956.t1; a = -2.78), was significantly depleted in adult haemolymph after immune challenge.

Table 14 Enriched or depleted proteins in *C. floridanus* larval and/or adult worker haemolymph upon immune challenge. Enrichment and depletion are given as log2(x-fold changes). The proteins listed have either known immune functions or have been shown to be regulated upon immune challenge in previous works. Each protein within the list was non-significantly regulated. (n.r. = not regulated).

acc # plus annotation	log2(x-fold change) adults injected	log2(x-fold change) larvae injected
Cflo_N_g8446.t1|Serine protease persephone	2.04	1.48
Cflo_N_g5742.t1|Beta-1,3-glucan-binding protein (GNBP)	1.82	n.r.
Cflo_N_g7945.t1|Serine protease snake 2	1.82	1.88
Cflo_N_g8262.t1|Vitellogenin-3 (Fragment)	1.46	-1.24
Cflo_N_g10889.t1|Hemolymph lipopolysaccharide-binding protein	1.35	1.73
Cflo_N_g102.t2|PGRP-SC2	1.17	n.r.
Cflo_N_g12735.t1|Protein Spätzle	1.04	n.r.
Cflo_N_g11181.t1|Hemocytin	1.01	n.r.
Cflo_N_g7714.t1|Transferrin	-1.26	1.03
Cflo_N_g11205.t1|Hexamerin	-1.31	-1.21
Cflo_N_g10612.t1|Hexamerin	-3.34	n.r.
Cflo_N_g10613.t1|Hexamerin	n.r.	-1.51
Cflo_N_g8526.t1|PGRP2 (=PGRP-SA)	n.r.	1.56
Cflo_N_g4036.t1|Lysozyme c-1	n.r.	1.55
Cflo_N_g10406.t1|Soma ferritin	n.r.	1.32
Cflo_N_g1203.t1|Ferritin subunit	n.r.	-1.21

As mentioned before, intracellular proteins were detected in the *C. floridanus* haemolymph and several of these proteins were highly enriched upon immune challenge of larvae and adults, e.g. the Glycine cleavage system H protein (Cflo_N_g938.t1; a = 4.70; l = 4.54), the Enoyl-CoA hydratase domain-containing protein 3 (Cflo_N_g9457.t1; a = 4.11; l = 6.66), the Myosin light chain (Cflo_N_g5638.t1; a = 2.94; l = 2.52) and the Tropomyosin (Cflo_N_g8574.t1; a = 4.91; l = 3.08). However, the abundance of cellular proteins might be a result of injury mediated rupture of cells leading to the release of cellular proteins when pricking the animals. Also, cellular mechanisms of the immune response, e.g. phagocytosis, might lead to the release of cellular components and to their enrichment upon immune challenge.

Recently, it was reported that the expression of several genes encoding proteins participating in metabolism is downregulated in *C. floridanus* after immune challenge (Gupta et al., 2015). Various proteins involved in metabolism are depleted in haemolymph samples of either immune-challenged workers or larvae or both developmental stages. The two proteins Zinc metalloprotease ZmpB (Cflo_N_g6298.t1; a = -2.57; l = -3.30) and beta-Galactosidase (Cflo_N_g5481.t1; a = -2.68; l = -2.79) were significantly depleted in both samples (Table 11), and a Hexamerin (Cflo_N_g11205.t1; a

= -1.31; l = -1.21) was non-significantly depleted in the haemolymph (Table 14), whereas other proteins such as Catalase (Cflo_N_g7519.t1; a = -2.99), Phosphoglycerate kinase (Cflo_N_g10398.t1; a = -3.92), the Adenosine kinase 2 (Cflo_N_g13339.t1; a = -4.27), Transferrin (Cflo_N_g7714.t1; a = -1.26) and another Hexamerin protein (Cflo_N_g10612.t1; a = -3.34) were significantly or non-significantly depleted only in the haemolymph of immune-challenged workers (Table 12). However, in larval haemolymph samples a Chymotrypsin inhibitor (Cflo_N_g10681.t1; l = 6.91) and a Chymotrypsin inhibitor-like protein (Cflo_N_g10682.t1; l = 2.44) were significantly enriched and several metabolically active proteins were significantly depleted, for example Trypsin-1 (Cflo_N_g4442.t1; l = -4.51), Maltase-1 (Cflo_N_g8234.t1; l = -3.86), alpha-Glucosidase (Cflo_N_g1453.t1; l = -2.75), a Hexamerin protein (Cflo_N_g10613.t1; l = -1.51) or Chymotrypsin-1 (Cflo_N_g7296.t1; l = -1.89) (Table 13 and Table 14). The results suggest the down-modulation of various metabolic activities following an immune challenge in *C. floridanus* and are therefore in concert with results from the immuno-transcriptome above (chapter 4.1.4). The differential expression analysis showed the downregulation of several genes encoding proteins (Gupta et al., 2015) which were also depleted in the haemolymph of immune-challenged animals, e.g. Trypsin-1 (Cflo_N_g4442.t1; a = -4.51), Hexamerin (Cflo_N_g10613.t1; l = -1.51), Maltase-1 (Cflo_N_g8234.t1; l = -3.86), Zinc metalloprotease ZmpB (Cflo_N_g6298.t1; a = -2.57; l = -3.30), and beta-Galactosidase (Cflo_N_g5481.t1; a = -2.68; l = -2.79). However, the expression of some genes was not regulated upon infection (Gupta et al., 2015), while the respectively encoded proteins were depleted in the haemolymph, for example Catalase (Cflo_N_g7519.t1; a = -2.99) and a Hexamerin (Cflo_N_g11205.t1; a = -1.31; l = -1.21). Also, while the expression levels of some genes increased after infection (Gupta et al., 2015), the gene products were depleted in the haemolymph, e.g. Transferrin (Cflo_N_g7714.t1; a = -1.26) and Chymotrypsin-1 (Cflo_N_g7296.t1; l = -1.89).

4.2.3 Haemolymph peptides in *C. floridanus*

Naturally occurring peptides or polypeptides are short stretches of amino acids or small proteins of low molecular weight. This highly diverse group comprises biologically active molecules such as hormones, neuropeptides, cytokines, growth factors and antimicrobial peptides. They have been shown to be involved in biological processes such as cell-cell communication, neuromodulation, and the regulation of physiological processes such as homeostasis and even behaviour (Mylonakis et al., 2016; Schmitt et al., 2015; Soloviev and Finch, 2006). In *Drosophila*, neuropeptides and peptide hormones are produced via the regulated secretory pathway (Pauls et al., 2014). These peptides are expressed as larger preproproteins which contain an N-terminal signal peptide. Therefore, the precursors are directed to the lumen of the rough endoplasmatic reticulum where the signal peptide is removed. The resulting proprotein can be modified, e.g. glycosylated, and is packed within dense-core vesicles within the trans-Golgi (Pauls et al., 2014). Within those vesicles, the proproteins are further processed and modified by a set of specific enzymes, e.g. the prohormone convertase 2 (AMON), the carboxypeptidase D (SILVER) and the amidating enzymes PAL 1/2 and PHM (Fricker, 2012; Pauls et

al., 2014; Veenstra, 2000). However, it is known that the precise *in silico* prediction of neuropeptides is difficult due to various modifications. Even if a prepropeptide is expressed in a certain insect stage, it is not clear which bioactive peptides are produced (Schmitt et al., 2015). Though the genes coding for homologs of the required enzymes of the regulated secretory pathway are present in the *C. floridanus* genome (Bonasio et al., 2010; Gupta et al., 2015), active peptides have to be identified directly. Recently, the neuropeptidome of adult *C. floridanus* was biochemically identified by combining different mass spectrometric methods (Schmitt et al., 2015). To gain further insights into the immune-related peptidome in *C. floridanus,* peptides were isolated and identified from haemolymph samples of immune-challenged and untreated larvae or adult animals with mass spectrometry (chapter 3.7).

4.2.3.1 Peptides in the haemolymph of larvae

In total 1,717 peptides were identified in the haemolymph of *C. floridanus* larvae. Only peptides which showed differential regulation after immune challenge (708) were chosen for further analyses whereas the 1,009 unregulated peptides were excluded. Signal peptides of the respective proteins to which the identified peptides belong were identified using the SignalP 4.1 Server (Nielsen et al., 1997) and proteins were further processed *in silico* using the NeuroPred online tool (Southey et al., 2006). Following this workflow, 567 regulated peptides of the larval haemolymph were assigned to 39 proteins with predicted peptides according to the NeuroPred online tool. Remarkably, 484 regulated peptides derived from the AMP hymenoptaecin. All of these peptides exhibit different amino acid sequences or show at least different posttranslational modifications such as oxidation (digital Appendix, Table S3). More than 50% of the other 38 proteins are annotated as putative uncharacterised proteins (digital Appendix, Table S3).

As described in Methods chapter 3.7.3 only peptides with intensity values for at least 4 time points were considered and the log2-fold expression changes were calculated against the t = 0 time point. The temporal peptide profiles were clustered according to their kinetics following an immune challenge utilising fuzzy c-means clustering (performed by Jens Vanselow; (Futschik and Carlisle, 2005)). Cluster 0 contains all unregulated peptides and the 708 regulated peptides of the larval haemolymph were categorised into four different regulatory clusters (Figure 22). Cluster 1 and cluster 3 comprise 427 peptides, which were upregulated at later time-points after an infection (24 hours till 72 hours post injection with an at least log2-fold expression change of 1). Peptides belonging to cluster 3 showed highest expression levels at 24 hpi, their expression levels were downregulated at 48 hpi and again upregulated at 72 hpi. In contrast, the clusters 2 and 4 include only 281 peptides, which were already upregulated at early time points after infection. Peptides belonging to cluster 2 were already upregulated at 6 hpi, whereas peptides of cluster 4 reached their highest expression levels at 12 hpi. Peptides from both, cluster 2 and 4, were depleted in larval haemolymph at later time points after immune challenge.
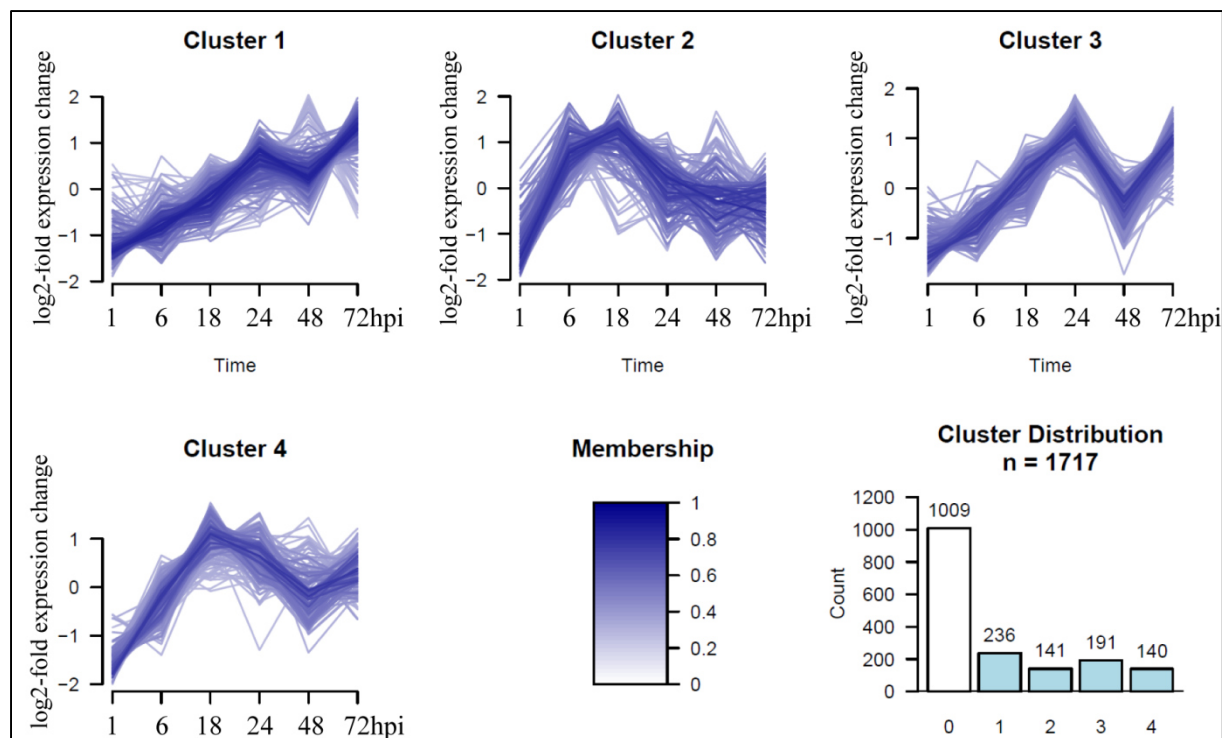
Figure 22 Cluster formation of regulated peptides in haemolymph of *C. floridanus* larvae. 1,717 identified peptides in the haemolymph of *C. floridanus* larvae (Lsupern0235) are pooled into 5 different clusters describing the kinetics of peptide regulation after an immune challenge. Cluster 0 includes unregulated peptides. Cluster 1 and 3 include peptides which were upregulated at later time points after an immune challenge, whereas cluster 2 and 4 comprise peptides upregulated at early time points after an immune challenge. The membership of a peptide to its cluster is indicated by blue colour with dark blue highlighting a hundred percent membership.

Noticeably, the identified peptides in the samples have molecular weights of less than 3 kDa, whereas the *in silico* peptides have significantly higher molecular weights. This observation is in concert with the previously described discrepancies between *in silico* predicted peptides and the biochemically identification of peptides (Schmitt et al., 2015). Furthermore, predicted peptides with molecular weights > 10 kDa were considered to be proteins and very likely to precipitate during the separation of peptides and proteins as described above (chapter 3.7.1). However, one identified peptide (#9221) comprises the exact sequence and same size of about 1,257 Dalton as the predicted peptide belonging to the protein previously annotated as Vitellogenin (EAG_11463) (Table 15). The peptide derives from the C-terminus of the Vitellogenin protein.

Haemolymph peptides derived from the other 38 proteins mostly consist of only a part of the *in silico* predicted peptides. For example two peptides with molecular weights < 2 kDa (# 992: 1.40 kDa; # 7016: 1.15 kDa) were identified and derived from the putative uncharacterized protein (EAG_04523) which has two *in silico* predicted peptides with a molecular weights of 5.09 kDa and 4.68 kDa (Table 15; digital Appendix, Table S3). Interestingly, four haemolymph peptides (# 10300: 1.36 kDa; #9592: 1.26 kDa; # 6793: 1.33 kDa; # 9788: 1.16 kDa) were identified and derived from the defensin-1 (GenBank: JN989495.1 (Ratzka et al., 2012a)) (Table 15; digital Appendix, Table S3). The peptides belong to cluster 4 and were enriched at early time points after immune challenge. The identification of these peptides from defensin-1 supports the transcriptome and proteome data

described in chapters 4.1.4 and 4.2.2 (Table 10 and Table 13). The expression of the *defensin-1* gene was upregulated upon immune challenge and the AMP was highly enriched within the haemolymph protein samples of immune-challenged larvae and adults. The peptides identified in the samples consist of parts of one predicted peptide which derives from the N-terminus of the defensin-1 (4.9 kDa). However, the 4.04 kDa peptide, consisting of 40 amino acids and deriving from the C-terminus of the precursor molecule, was predicted to be the functional mature defensin-1 peptide in a previous analysis (Ratzka et al., 2012a). Therefore, it is not clear whether the defensin-1 peptides identified in haemolymph peptide samples indicate depletion of the full-length N-terminal defensin-1 peptide at early time points after immune challenge, or if these peptides have bioactivity in addition to the predicted functional defensin-1 peptide. The latter was solely identified in the haemolymph protein samples which could be a consequence of its molecular weight of > 4 kDa. As mentioned before, only peptides with molecular weights of < 3 kDa were detected in the peptide samples.

Table 15 Regulated peptides within the larval haemolymph. Exemplarily, the table lists peptides derived from three *C. floridanus* proteins and regulated upon immune challenge, their regulatory clusters, the peptide sequences and molecular weights. Additionally amino acid sequences for the *in silico* predicted peptides of the proteins are given and the sequences of the haemolymph (hl) peptides were aligned to display matching sequences (bold printing).

	Cluster	peptide ID	peptide sequence	peptide mass [Dalton]
Vitellogenin OS=*Camponotus floridanus* GN=EAG_11463				
in silico			**TLSNIYTVPHL**	1257.45
hl	1	9221	**TLSNIYTVPHL**	1256.67
Putative uncharacterized protein OS=*Camponotus floridanus* GN=EAG_04523				
in silico			NSQYAPIDIANVN**DIEKIFTDQLPS**MLKDVFNSFVQDKNEFSKK	5092.71
hl	3	992	**DIEKIFTDQLPS**	1404.71
in silico			KLEILEKILRDSEIRDNKWKTKGIER**QIERIQQRT**RR	4677.44
hl	4	7016	**QIERIQQRT**	1153.62
Defensin-1 GN=JN989495.1				
in silico			FPTEELELEEN**VTVESPDFLILK**KDKSL**QETPIKEHNRT**RR	4910.52
hl	4	10300	**VTVESPDFLILK**	1359.76
hl	4	9592	**TVESPDFLILK**	1260.69
hl	4	6793	**QETPIKEHNRT**	1334.65
hl	4	9788	**VESPDFLILK**	1159.64
in silico			ATCDLLSGFGVNHSACAAHCILRGKTGGRCNSNAVCVCRA	4037.66

As mentioned above, 484 regulated peptides with different amino acid sequences or posttranslational modifications were detected in the haemolymph of *C. floridanus* larvae and assigned to the predicted peptides (hymenoptaecin-like peptide or hymenoptaecin repeat domain peptide) of the AMP hymenoptaecin. The peptides are found in all four regulatory clusters, and peptides within each cluster consist of major parts of the *in silico* predicted hymenoptaecin peptides (digital Appendix, Table S3). 162 regulated peptides belong to cluster 1 and 3 and were upregulated at late time points after

infection. The majority of the peptides (322) were enriched early (6 or 12 h) after immune challenge (digital Appendix, Table S3). However, the predicted functional peptides encoded by the *hymenoptaecin* gene have molecular weights above 10 kDa. Therefore, it is more likely that the supposedly functional peptides are precipitated with the haemolymph proteins. As described before, hymenoptaecin (Cflo_N_g14777.t1; a = 2.45; l = 4.70) was not only detected in the ants' haemolymph protein samples but also highly enriched after immune challenge in larvae and workers (chapter 4.2.2).

4.2.3.2 Peptides in haemolymph of workers

In total, 2,181 peptides were identified in two replicate haemolymph samples of *C. floridanus* workers. The majority of peptides (1,179) in worker haemolymph showed no differences in expression after an immune challenge. These peptides were excluded from further analyses. 342 out of 1002 differentially regulated peptides in worker haemolymph samples derived from 13 proteins with predicted peptides. Similar to the larval haemolymph sample more than 50% of these proteins (7/13) are annotated as putative uncharacterised proteins (digital Appendix, Table S4). Based on their characteristic kinetics, the 1002 regulated peptides were clustered into five groups (cluster 1 to cluster 5) (Figure 23). Cluster 1 and cluster 3 include 386 peptides, which are upregulated at later time-points after an immune challenge, with cluster 3 containing the peptides upregulated at very late time points (48-72 hpi). In contrast, cluster 2 includes 207 peptides which are upregulated already at 6 hpi and showed a continuously high expression at all later time points (till 72 h). In addition, 409 peptides belonging to cluster 4 and 5 were upregulated at 6 hpi, but were depleted in the haemolymph of workers at later time points (Figure 23).

Similar to larval haemolymph samples, the majority of haemolymph peptides in workers were assigned to hymenoptaecin (316 peptides; digital Appendix, Table S4). The cluster distribution of these peptides, which show different amino acid sequences, indicates an overall high abundance of hymenoptaecin derived peptides in the haemolymph at late time points after immune challenge. In contrast to the observations in larval haemolymph, most of the worker haemolymph peptides belong to cluster 1 and 3 and were therefore enriched at later time points after immune challenge. Only 54 peptides were distinctly upregulated at 6 hpi and continuously depleted at later time points which might suggest a role of these peptides during the early stages of the immune response. However, each peptide itself consists of only a short part of the amino acid sequences of the predicted hymenoptaecin peptides derived from the hymenoptaecin-like domain or the hymenoptaecin repeat domains. It is not clear whether these peptides are functionally active, or if they are the result of the degradation of the supposedly functional hymenoptaecin peptides which were detected in the haemolymph protein samples of adult workers (chapter 4.2.1).
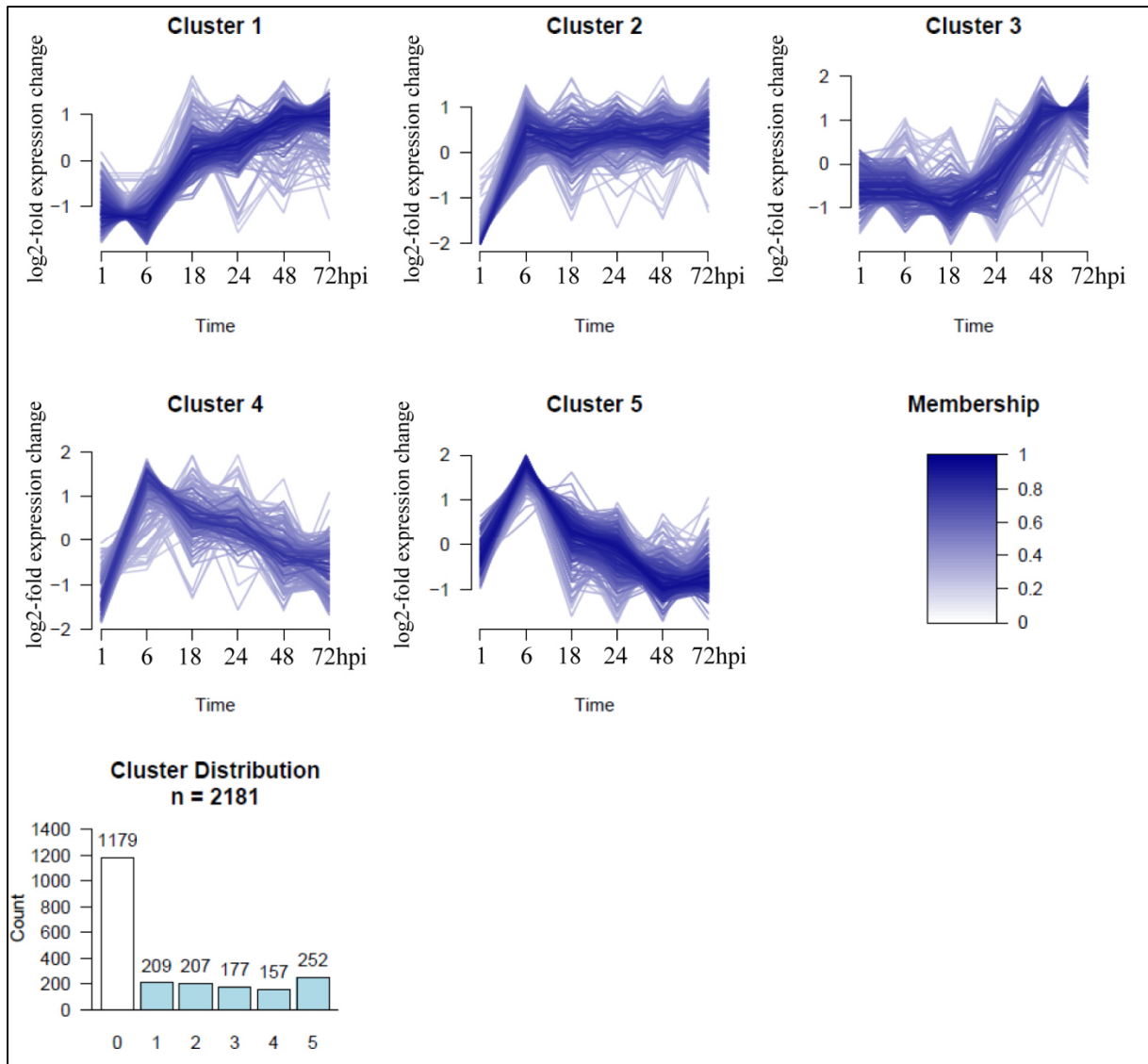
Figure 23 Cluster formation of regulated peptides in the haemolymph of immune-challenged *C. floridanus* workers. 2,181 identified peptides in the haemolymph of *C. floridanus* adult workers (two replicates Supern0235 and Supern0178) are pooled into 6 different clusters describing the kinetics of peptide regulation after an immune challenge. Cluster 0 includes not-regulated peptides (1,179). Cluster 1 and 3 include peptides which are upregulated at later time points after an immune challenge with cluster 3 including peptides only upregulated after 48 h and 72 h. Cluster 2 holds peptides upregulated early (6 h) and showing constantly high expression during later time points. Cluster 4 and 5 include peptides upregulated early after an immune challenge (6 h) but with lower expression at later time points. The membership of a peptide to its cluster is indicated by blue colour with dark blue highlighting a hundred per cent membership.

Five regulated peptides with different amino acid sequences belonging to defensin-2 (Genbank: JQ693412.1 (Ratzka et al., 2012a)) were identified in the adult haemolymph peptide samples (# 3711: 2.11 kDa); # 3712: 2.13 kDa; # 5705: 2.43 kDa; # 9076: # 9534: 1.61 kDa). The five peptides were highly upregulated at 6 hpi but depleted afterwards (Table 16 ; digital Appendix, Table S4). Thus, these peptides belong to cluster 5 in either both or just one of the two replicates. However, none of the peptides consists of the complete amino acid sequence of the predicted functional defensin-2 peptide with a molecular weight of 4.88 kDa (Ratzka et al., 2012a) but comprise parts of another cleavage product of the defensin-2 precursor (4.15 kDa). Peptides comprising short parts of the *in silico* predicted functional peptide or the complete supposedly functional defensin-2 peptide were neither

detected in the peptide nor in the protein samples of larval or adult haemolymph (chapter 4.2.1 and 4.2.2). Additionally, the expression of the *defensin-2* gene was downregulated upon immune challenge of the ants (Gupta et al., 2015).

Table 16 Regulated peptides of the defensin-2 in the worker haemolymph. The table lists peptides derived from the *C. floridanus* defensin-2 and regulated upon immune challenge, their regulatory clusters, the peptide sequences and molecular weights. Additionally amino acid sequences for the *in silico* predicted peptides of the proteins are given and the sequences of the haemolymph (hl) peptides were aligned to display matching sequences (bold printing).

	Cluster	peptide ID	peptide sequence	peptide mass [Dalton]
			Defensin prepropeptide OS=*Camponotus floridanus* (Defensin-2 (JQ693412.1))	
in silico			NTLSAVYDGPTYE**LTTIDEPQYDEMASNLSPIRH**RR	4154.53
hl	5	3711	**IDEPQYDEMASNLSPIRH**	2113.97
hl	5	3712	**IDEPQYDEMASNLSPIRH (oxidation)**	2129.96
hl	5	5705	**LTTIDEPQYDEMASNLSPIRH**	2429.15
hl	5	9076	**TIDEPQYDEMASNLSPIRH**	2215.02
hl	5	9534	**TTIDEPQYDEMASN**	1612.65
in silico			VTCDLLSWQSQWLTINHSACAAKCLVQRRRGGRCRDGICVCRN	4878.68

4.3 The detection of hymenoptaecin peptides in the haemolymph from *C. floridanus* with a newly generated antiserum

One of the fastest and specific methods to identify a protein within a sample is the detection of proteins via antibodies in Western blots. In order to quickly identify the antimicrobial peptide hymenoptaecin in haemolymph samples of immune challenged *C. floridanus* an anti-hymenoptaecin antiserum was produced (chapter 3.9.). The recombinant AMP was expressed in *E. coli* Rosetta 2(DE3)pRARE2/pET15b_*hrep*10 and purified under denaturing conditions which was sufficient for the following antiserum production. The antigen hymenoptaecin was injected into a rabbit and several vials of antiserum were obtained at different time points after injection of the antigen, and quality and specificity of the antiserum were tested.

In a first analysis the anti-hymenoptaecin antiserum was tested against the purified hymenoptaecin peptide sample which was used for antibody production. The antiserum detected the 6xHis-hymenoptaecin fusion protein which has a molecular weight of about 12 kDa (Figure 24). Additionally, the antiserum detected several additional proteins which remained as impurities within the 6xHis-hymenoptaecin sample even after extensive purification.
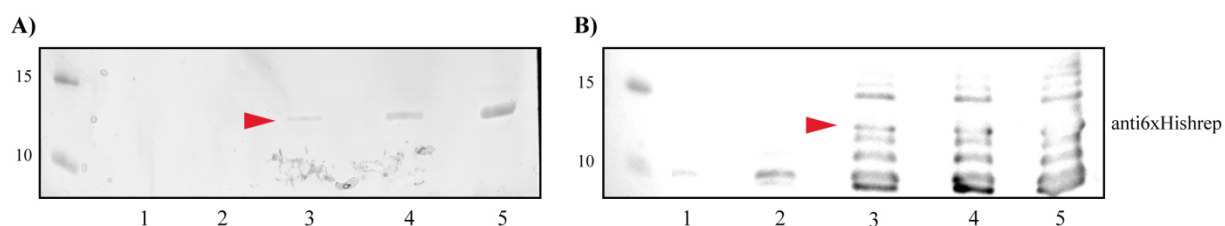
Figure 24 Detection of 6xHis-hymenoptaecin fusion proteins. Samples of purified 6xHis-hymenoptaecin fusion proteins were analysed by **(A)** SDS-PAGE and **(B)** Western blotting. Different amounts of the fusion proteins were loaded on the gels: 1) 0.06 µg; 2) 0.12 µg; 3) 0.6 µg; 4) 1.2 µg and 5) 3 µg. The antiserum produced in this work was used for detection of the fusion proteins. The fusion proteins were detected in samples with a total protein amount higher than 0.12 µg (marked with an arrow) either by Coomassie staining **(A)** or via the antiserum **(B)**.

Furthermore, the capability of the antiserum to detect hymenoptaecin peptides in haemolymph samples from *C. floridanus* was analysed. For this purpose, workers taken from one colony (*C. floridanus* C90) were either not treated (control) or injected with a mixture of heat-killed Gram-positive and Gram-negative bacteria. Haemolymph samples were taken in triplicates at different time points after infection and analysed via Western blot. The hymenoptaecin peptides were detected by the antiserum and an upregulation of the hymenoptaecin within the haemolymph at 18 hpi and at later time points was observed. These results are in accordance with the proteome and the peptidome data, which showed enrichment of hymenoptaecin in the worker haemolymph upon immune challenge (Table 13) and most of the identified hymenoptaecin peptides were enriched in haemolymph peptide samples of workers later than 18 hpi (digital Appendix, Table S4). In one replicate the antiserum detected the putative AMP already at 6 hpi (Figure 25 A) which is in accordance with the identification of several peptides which were enriched in worker haemolymph at earlier time points

after immune challenge (chapter, 4.2.3; digital Appendix, Table S4). Thus, the results suggest a general immune-induced upregulation of the expression of the hymenoptaecin and the release of the peptides with molecular weights between 10 and 12 kDa into the haemolymph, although the expression kinetics may vary between individuals or groups of animals even within one colony.
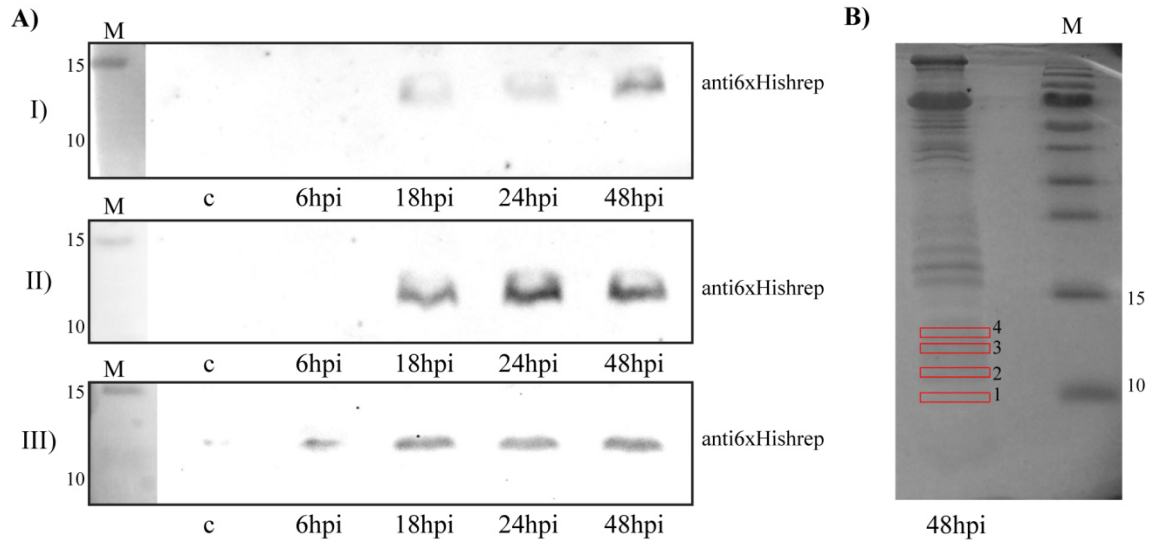
Figure 25 Detection of hymenoptaecin peptides in haemolymph samples of *C. floridanus*. **(A)** Haemolymph samples taken from *C. floridanus* workers (three replicates I – III) from colony C90) at different time points after an immune challenge were analysed further by SDS-PAGE and Western blotting. Hymenoptaecin peptides were detected using the produced anti-hymenoptaecin antiserum of this work. **(B)** Detection of hymenoptaecin peptides via SDS-PAGE and Coomassie staining. Coomassie-stained protein bands (1-4) containing proteins with molecular weights of 10-15 kDa were cut out and further analysed with mass spectrometry.

In addition to the antibody based detection of the hymenoptaecin peptides, protein bands with the molecular weight of 10-15 kDa were cut out of a SDS gel and proteins were identified via mass spectrometry (Figure 25 B). Proteins of band 1 could not be identified. The proteins in bands 2 to 4 were identified as hymenoptaecin peptides encoded by the hymenoptaecin repeat domains (HD) (Table 25) (Ratzka et al., 2012a). Interestingly, although these hymenoptaecin peptides have a predicted molecular weight of 10.4 kDa to 10.7 kDa the bands run at a height corresponding to 11-13 kDa. The different electrophoretic mobility of the peptides on the SDS gel might be the result of posttranslational modifications such as phosphorylation. Altogether, the mass spectrometry results confirmed that the produced anti-hymenoptaecin antiserum specifically detects hymenoptaecin peptides in haemolymph samples of immune-challenged *C. floridanus* workers.

4.4 Recombinant expression, purification and functional analysis of hymenoptaecin derived peptides from *C. floridanus*

4.4.1 Expression of hymenoptaecin derived peptides in insect cells

The hymenoptaecin gene of *C. floridanus* encodes a multipeptide precursor protein containing a signal peptide and a propeptide which is followed by a single hymenoptaecin-like domain and six repeated hymenoptaecin domains. The sequences coding for seven putative hymenoptaecin peptides are separated from each other by two probable processing sites "EAEP" and "RR" with the latter alternatively comprising a combination of lysine and arginine ("KR") (Figure 26 and Figure 27). The expression of the *hymenoptaecin* gene is upregulated upon infection (Gupta et al., 2015; Ratzka et al., 2011) and it was suggested that the processing of the multipeptide precursor would result in the release of seven putative hymenoptaecin peptides (Ratzka et al., 2012a). Whereas the putative mature hymenoptaecin peptides are 97 amino acids long and start with an N-terminal glutamine residue (Q) (Figure 27 A), the mature hymenoptaecin-like peptide is 108 aa long and its amino acid sequence begins with the motif "DAEK" followed by a glycine (G). Additionally, the mature hymenoptaecin-like peptide comprises an N-terminal insertion of 14 aa (Figure 26 A). Due to these structural differences, it can be assumed that the hymenoptaecin derived peptides also vary in their function. For this purpose, DNA fragments encoding a hymenoptaecin-like peptide (hld) and one hymenoptaecin peptide encoded by a repeat domain (hrep) were cloned into two different plasmid vectors.

In this work, a baculovirus expression vector system (BEVS) in insect cells was established for the production of hymenoptaecin peptides. The expression system based on the multigene baculoviral vector MultiBac enables the production of high amounts of multiple soluble proteins which exhibit post-translational modifications (Fitzgerald et al., 2006). Therefore, two derivates of the pFBDM4 plasmid containing the sequences for two different affinity tags for protein purification in their multiple cloning sites (MCS1), a Glutathione-S-transferase (GST)-tag and a 6xHis-tag were used. In both plasmids, pFBDM4_mutGST[TEV](MCS1)_*egfp*(MCS2) and pFBDM4_6xHis[TEV](MCS1)_*egfp*(MCS2), sequences encoding the affinity tags are linked to the gene of interest via a sequence encoding the specific cleavage side ("ENLYFQG") for the Tobacco Etch Virus nuclear-inclusion-a endopeptidase (TEV). The constructs produced in this work are composed as follows: the fragment encoding the hymenoptaecin-like peptide is linked via the TEV cleavage site to the GST-tag (Figure 26 C), whereas the fragment encoding the hymenoptaecin peptide is linked via "ENLYFQG" to a 6xHis-tag (Figure 27 C). The integration of the gene fragments coding for hld and hrep into the multiple cloning sites of the pFBDM4 derivates yielded in the transfer plasmids pFBDM4_mutGST[TEV]*hld*(MCS1)_*egfp*(MCS2) and pFBDM4_mut6xHis[TEV]*hrep*(MCS1)_*egfp*(MCS2). The genes encoding the respective fusion proteins were under the control of the viral *polh* promotor and the reporter gene coding for EGFP was under the control of the viral promotor *p10*. Each transfer plasmid was used for the integration of the DNA fragments coding for the

reporter gene and the fusion proteins into the attTn transposition site of the baculoviral genome (Multibac) of the *Autographa californica* Multicapsid Nuclear Polyhedrosis Virus (AcMNPV) and recombinant bacmid DNA was purified. Afterwards, insect cells *Sf*21were transfected with the purified bacmid DNA for virus production as described in chapter 3.10.1.
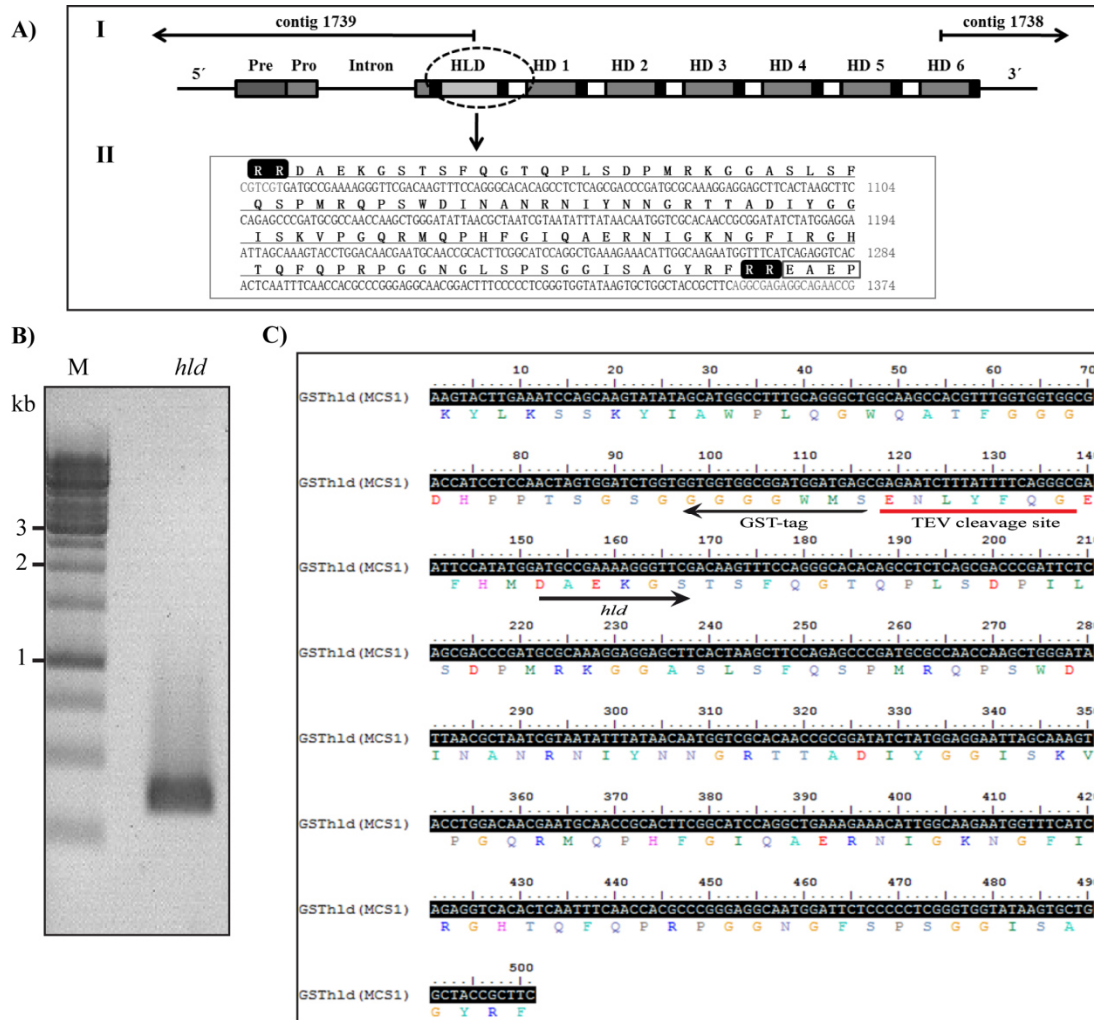
Figure 26 The GST-hymenoptaecin-like peptide fusion protein. **(A)** The schematic structure of the *hymenoptaecin* gene. I) The gene encodes a multipeptide precursor protein consisting of a signal-sequence (Pre) and a pro-sequence (Pro), which is followed by a hymenoptaecin-like domain (HLD) and six repeated hymenoptaecin domains (HD 1–6). The hymenoptaecin domains are flanked by the two putative processing sites "EAEP" (white boxes) and "RR" (black boxes). II) The nucleotide and deduced amino acid sequence of the hymenoptaecin-like peptide are shown (Ratzka et al., 2012a). **(B)** PCR amplification of the DNA fragments encoding the hymenoptaecin-like peptide. The PCR-products from cDNA were separated on a 1.2% agarose gel alongside molecular size markers (lane M, GeneRuler 1 kb DNA Ladder; kb = kilo base pairs) and analysed with HD Green staining (INTAS Science Imaging Instruments GmbH). The major band (lane *hld*) corresponds to the 336 bp hymenoptaecin-like peptide DNA fragment. **(C)** Composition of the GST-hymenoptaecin-like peptide fusion protein. The DNA fragment encoding the hymenoptaecin-like peptide (*hld*) is linked via the TEV cleavage site to the GST-tag (partially depicted). The sequence encoding the fusion protein is incorporated into the multiple cloning site 1 (MCS1) of the plasmid pFBDM4_mutGST[TEV]*hld*(MCS1)_*egfp*(MCS2). The complete sequence encodes a fusion protein GSThld with a molecular weight of about 40 kDa.
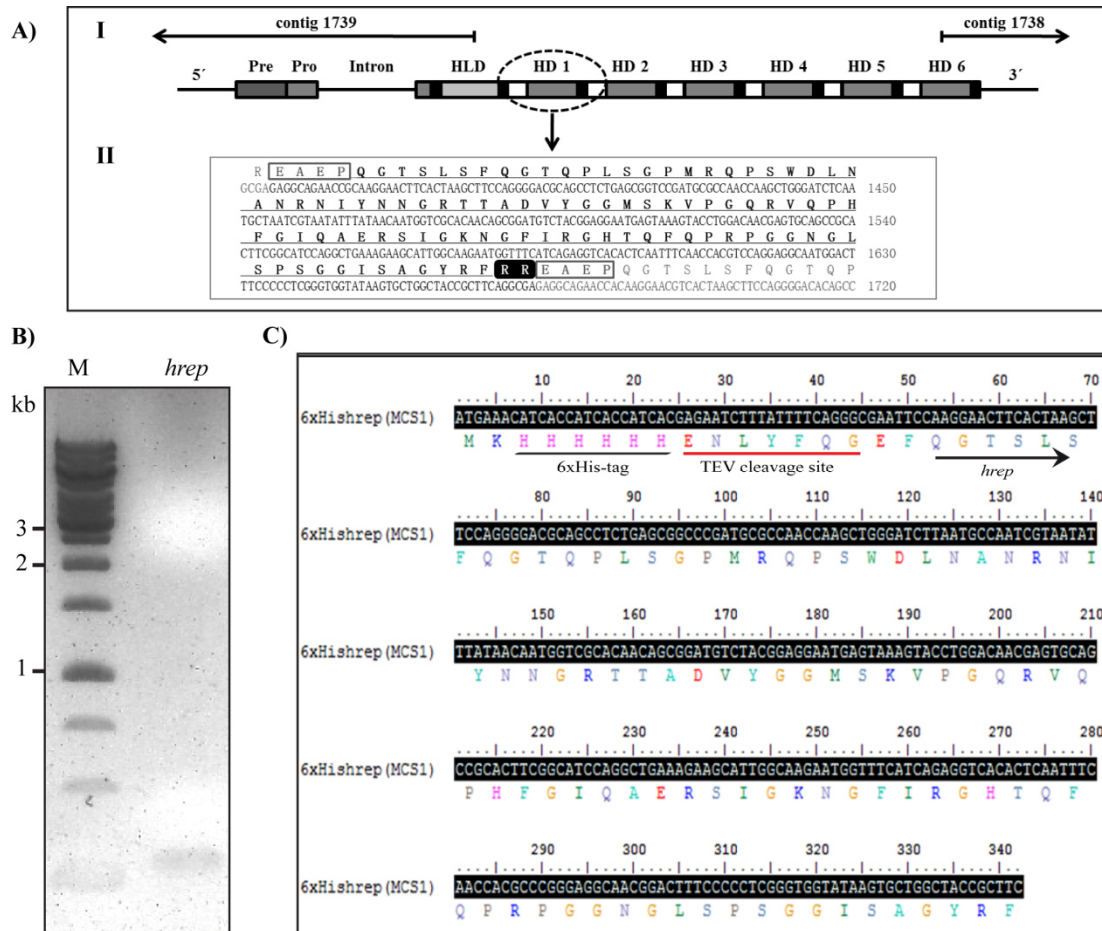
Figure 27 The 6xHis-hymenoptaecin peptide fusion protein. **(A)** The schematic structure of the *hymenoptaecin* gene I) which encodes a multipeptide precursor protein consisting of a signal-sequence (Pre) and a pro-sequence (Pro), followed by a hymenoptaecin-like domain (HLD) and six repeated hymenoptaecin domains (HD 1–6). The hymenoptaecin domains are flanked by the two putative processing sites "EAEP" (white boxes) and "RR" (black boxes). II) The nucleotide and deduced amino acid sequence of a hymenoptaecin domain encoded peptide are shown. **(B)** PCR amplification of the DNA fragments encoding the hymenoptaecin repeat domain derived peptide. The PCR-products from cDNA were separated on a 1.2% agarose gel alongside molecular size markers (lane M, GeneRuler 1 kb DNA Ladder; kb = kilo base pairs) and analysed with HD Green staining (INTAS Science Imaging Instruments GmbH). The major band (lane *hrep*) corresponds to the 291 bp hymenoptaecin domain DNA fragment. **(C)** Composition of the 6xHis-hymenoptaecin peptide fusion protein. The DNA fragment encoding the hymenoptaecin peptide (hrep) is linked via the TEV cleavage site to the 6xHis-tag. The sequence coding for the fusion protein is incorporated into the multiple cloning site 1 (MCS1) of the plasmid pFBDM4_6xHis[TEV]*hrep*(MCS1)_*egfp*(MCS2). The complete sequence encodes a fusion protein 6xHishrep with a molecular weight of about 13 kDa.

The transfection success was monitored utilising the availability of the non-invasive transfection marker EGFP which allows a rapid screening via fluorescence microscopy. Nine to ten days after transfection, EGFP positive fluorescence indicates the expression of genes under the control of viral promotors such as the reporter gene *egfp* and therefore the formation of budded viruses. For virus stock generation supernatants of beforehand transfected insect cell cultures were taken and fresh cultures of *Sf*21 cells were infected with the supernatants containing the recombinant baculoviruses. Using the end-point dilution method (chapter 3.10.1) the amount of viruses per millilitre within the virus stock (virus titre) was defined allowing the infection of insect cells with a defined MOI (multiplicity of infection) for protein production. Positively infected *Sf*21 cells expressed the green fluorescent EGFP after 2-3 days and the expression of the hymenoptaecin fusion proteins was further

verified by SDS-PAGE (Figure 28). Both constructs were suitable for the expression of the fusion proteins GSThld or 6xHishrep 3 days after infection of *Sf*21 cells with an MOI 0.5 for GSThld or MOI 1.0 for 6xHishrep. However, the expression levels of both proteins appeared to be low, as the respective protein bands have comparatively low intensity, whereas the EGFP was generally expressed in high amounts (Figure 28).
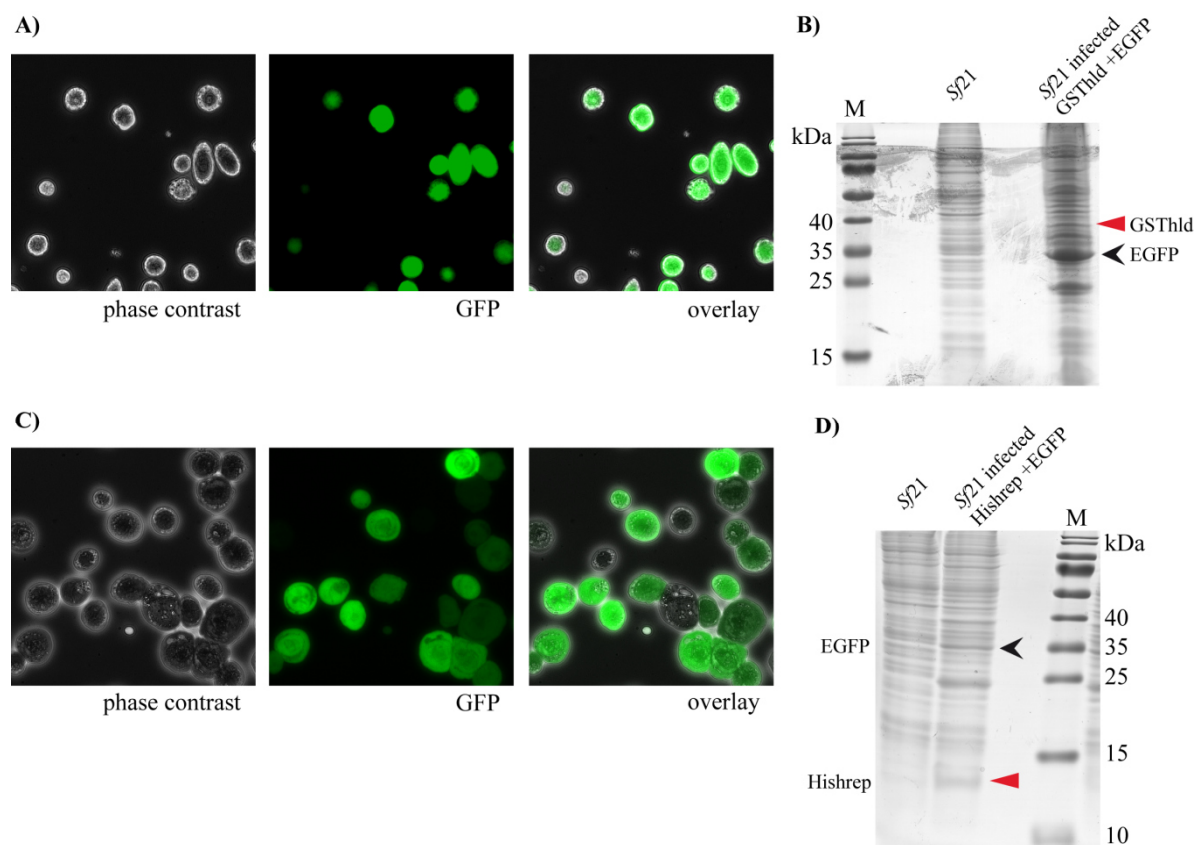
Figure 28 Expression of recombinant hymenoptaecin fusion proteins in insect cells. **(A)** The positive infection of *Sf*21 insect cells was confirmed via fluorescence microscopy. Baculovirus infected cells expressed the enhanced green fluorescent protein (EGFP). **(B)** SDS-PAGE of uninfected (lane 2) and baculovirus-infected *Sf*21 cells (lane 3) expressing GSThld (red arrow head) and EGFP (black arrow head) 3 days after infection with MOI 0.5. **(C)** Fluorescence microscopy of baculovirus infected *Sf*21 insect cells revealed positively infected cells due to the expression of EGFP. **(D)** SDS-PAGE of uninfected (lane 1) and baculovirus-infected *Sf*21 insect cells (lane 2), which expressed 6xHishrep (red arrow head) and EGFP (black arrow head) 3 days after infection with MOI 1.0.

4.4.2   Purification of GSThld and 6xHishrep from insect cells

The GST hymenoptaecin-like peptide (GSThld) fusion proteins were purified from *Sf*21 cells under different buffer conditions. In short, *Sf*21 cells from 50 ml cultures expressing EGFP and the fusion proteins were lysed via sonication in either Tris or PBS based buffers (chapter 2.6.4; Figure 29 A) lanes I-IV). Lysates were centrifuged to separate soluble from insoluble proteins. The soluble fractions were applied to glutathione beads for affinity purification. Unspecifically bound proteins were washed out, and specifically bound proteins were eluted via increased glutathione concentrations in the elution

buffer. During these purification steps, several difficulties occurred. Even though different Tris or PBS based buffers were used, only small amounts of the fusion proteins within the lysates were found to be soluble and therefore in the supernatants (Figure 29 A). However, the majority of proteins appeared to be insoluble and therefore in the pellets which was additionally shown by the purification of the GSThld fusion protein with the Cytobuster™ reagents. These reagents are suitable for the separation of soluble from insoluble proteins within a given sample of proteins expressed within insect cells. The analysis of the Cytobuster™ extraction showed that most of the GSThld fusion proteins expressed in insect cells were insoluble and therefore in the pellets (Figure 29 D). Furthermore, the abundance of the fusion proteins in the flow-through indicates that the binding of the GST-tagged fusion proteins to the glutathione resin appeared to be less than 100% (Figure 29 A), and only small amounts of the GSThld fusion proteins were eluted. However, the Tris based buffers (pH 7.5) were considered to be convenient for the purification of GSThld under native conditions which would be preferred with respect to maintaining functionality of the putative AMP throughout the purification procedure. Therefore, scale-up reactions were performed under these buffer conditions (Figure 29 A) V). Though fusion proteins were purified from an insect culture volume of 500 ml instead of 50 ml, the amount and concentration of GSThld in the elution fractions were very low. Additionally, a lot of fusion proteins remained bound to the glutathione beads and could not be removed applying mild elution conditions suitable for retrieval of native proteins. In order to avoid extensive re-folding procedures, the elution under harsh conditions was not taken into consideration. Overall, the purification of GSThld from *Sf*21 cells yielded only small amounts of fusion proteins and even scaling-up of expression and purification reactions did not improve these results. However, a test TEV protease cleavage of the GSThld fusion proteins purified from 50 ml insect culture was performed successfully (Figure 29 C). Measuring the amount of hymenoptaecin-like peptides after cleavage suggests a theoretical production rate of about 6 µg hymenoptaecin-like peptide per 50 ml insect cell culture (120µg/ litre).

Furthermore, only small amounts of the 6xHishrep fusion proteins could be purified using different standard buffer systems (chapter 2.6.4). As for the GSThld fusion protein, the majority of 6xHishrep remained in the insoluble fraction and was therefore not accessible for affinity purification under native conditions. The elution fractions did not contain detectable amounts of the fusion proteins (Figure 29 B). To verify the suspected low solubility of the 6xHishrep fusion proteins, the purification with Cytobuster™ reagents was performed and confirmed that the amount of soluble and therefore native and correctly folded 6xHishrep fusion proteins was very low (Figure 29 D).

It was considered that further purification steps to remove impurities from both fusion protein samples would result in the loss of more peptide with each purification step, e.g. due to irreversible binding of proteins to purification residues or beads. Thus, the low amount of fusion proteins expressed and purified per litre insect cell culture did not warrant the use of the cost-intensive and time-consuming

BEVS expression system for the production of high amounts of the putative AMPs (preferably 0.5-1.0 mg of each peptide for antimicrobial activity assays).
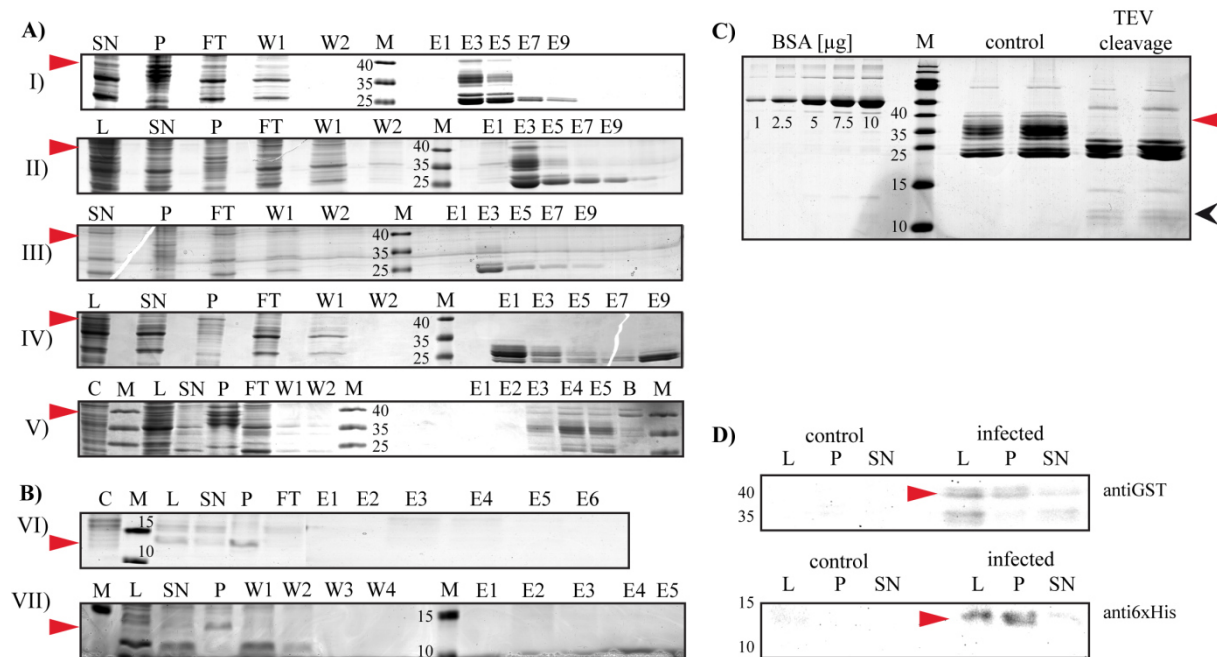

Figure 29 Purification of GSThld and 6xHishrep from *Sf*21 insect cells. **(A)** GSThld fusion proteins expressed in insect cells were purified using several buffers: I) GST Buffer 02; II) GST Buffer 03; III) GST Buffer 02, pH 8.0; IV) GST Buffer 01. V) GST Buffer 02 was used in a scale-up purification (buffer composition in chapter 2.6.4). The purification fractions were analysed via SDS-PAGE and the height corresponding to 40 kDa, which is the molecular weight of the fusion protein is indicated with red arrow heads. **(B)** 6xHishrep fusion proteins expressed in insect cells were purified using several buffers: VI) His Buffer 01; VII) His Buffer 03 (buffer composition in chapter 2.6.4). The purification fractions were analysed via SDS-PAGE and the height corresponding to 13 kDa, which is the molecular weight of the fusion protein is indicated with red arrow heads. The elution fractions appear 6xHishrep fusion protein free. **(C)** TEV cleavage of GSThld fusion proteins. Elution fractions of a GST purification containing about 190 µg of protein (fusion protein indicated with a red arrow head) were digested with the TEV protease. Different concentrations of BSA were used to determine the amount of hymenoptaecin-like peptides (black arrow head) via ImageJ. The digestion of fusion proteins isolated from 50 ml insect cell culture resulted in about 6 µg of the hymenoptaecin-like peptide. **(D)** Extraction of fusion proteins from insect cells via Cytobuster™. The Cytobuster™ reagents were used to isolate fusion proteins from insect cells, thereby defining whether the proteins were insoluble (P = pellet) or soluble (SN = supernatant). Most of the fusion proteins (GSThld and 6xHisHrep; indicated by red arrow heads) expressed in insect cells are insoluble. (L = lysate; FT = flow through; W = wash fraction; E = elution fraction; B = beads; M = molecular weight marker; molecular weights in kDa)

4.4.3 The expression of thioredoxin-S-His6 hymenoptaecin fusion proteins in *E. coli*

To avoid disadvantages such as high costs of the baculovirus expression vector system, a bacterial expression system was established in order to produce high amounts of the two hymenoptaecin derived peptides hld and hrep for antimicrobial activity assays. The protein expression in *E. coli* is a fast, cost-efficient and well-established standard system. The DNA fragments encoding the hymenoptaecin peptides were incorporated into the pET32 plasmid. Thus, the expression of the respective fusion proteins is regulated by an IPTG inducible lac-promotor. Furthermore, the recombinant plasmids were transformed into *E. coli* Rosetta 2(DE3) strains, which carry a plasmid

pRARE2 coding for seven tRNAs recognising the following rare codons: AUA (isoleucine), AGG, CGG und AGA (arginine), CUA (leucine), CCC (proline) and GGA (glycine) (Novy et al., 2001). Expression of proteins by these strains allows the so-called codon usage optimisation and the expression of proteins containing amino acids encoded by generally uncommonly used codons in *E. coli*. The integration of unspecific amino acids due to uncommon codons and pausing or early termination of translation are prevented (Kurland and Gallant, 1996; Sharp and Li, 1986).
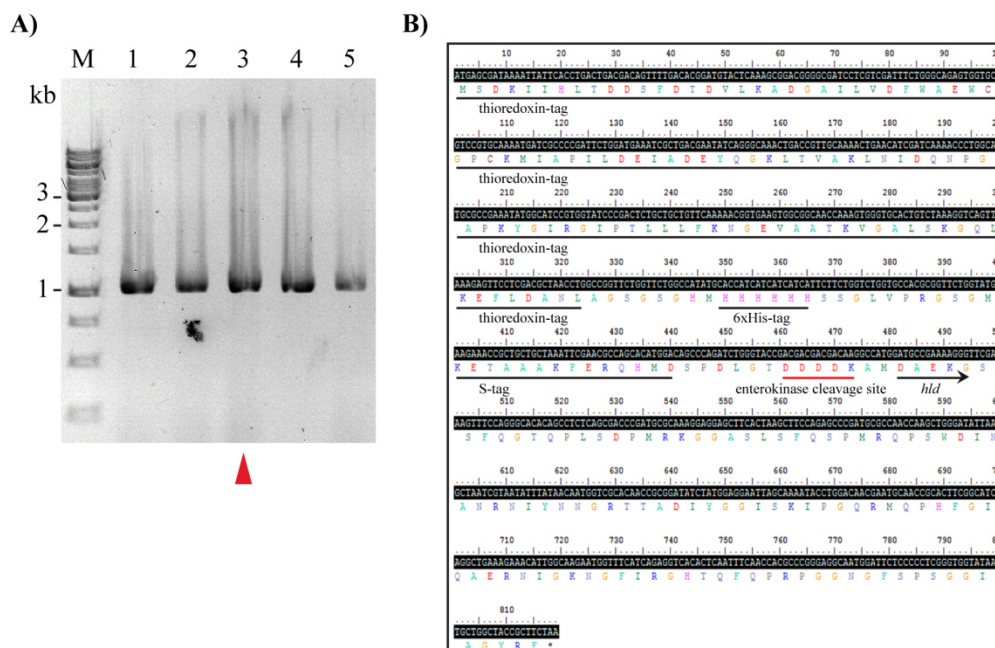
Figure 30 Composition of thioredoxin-S-His6 hymenoptaecin-like peptide fusion protein. **(A)** PCR amplification of the DNA fragments encoding the multiple cloning site (MCS) of pET32a. The fragments were amplified using the pET T7 oligonucleotides and plasmid DNA from *E. coli* Rosetta 2(DE3) clones transformed with recombinant pET32a was used as a template. The PCR products were separated on a 1.2% agarose gel alongside molecular size markers (lane M, GeneRuler 1 kb DNA Ladder; kb = kilo base pairs) and analysed with HD Green staining (INTAS Science Imaging Instruments GmbH). The major bands (lane 1-5) correspond to the about 1,000 bp MCS containing the sequence coding for the thioredoxin-S-His6 tagged hymenoptaecin-like peptide. The *E. coli* Rosetta 2(DE3)_pET32a_*hld* 3 clone was chosen for protein expression and purification (red arrow head). **(B)** The DNA fragment encoding the hymenoptaecin-like peptide (*hld*) was incorporated into the multiple cloning site of the plasmid pET-32a and is linked to the thioredoxin-S-His6-tag via the enterokinase cleavage site. The fusion tag consists of a thioredoxin-tag, followed by a 6xHis-tag and an S-tag. The complete sequence encodes a fusion protein thioredoxin-S-His6hld with a molecular weight of about 29.4 kDa. (* = Stop codon)

The fragments encoding hld and hrep were incorporated into the multiple cloning sites (MCS) of pET32a plasmids. The MCS follows a sequence coding for the thioredoxin-S-His6 tag, consisting of a thioredoxin-tag, a 6xHis-tag and an S-tag. While the full tag generally supports the solubility of expressed fusion proteins, affinity purification is enabled by the 6xHis-tag. Furthermore, the fusion proteins contain an enterokinase cleavage site for later removal of the thioredoxin-S-His6 tag resulting in recombinant hymenoptaecin peptides with an N-terminal extension of two amino acids (alanine and methionine) for the hymenoptaecin-like peptide (Figure 30) and three amino acids (alanine, methionine and alanine) for the hymenoptaecin peptide encoded by a repeat domain (Figure 31).
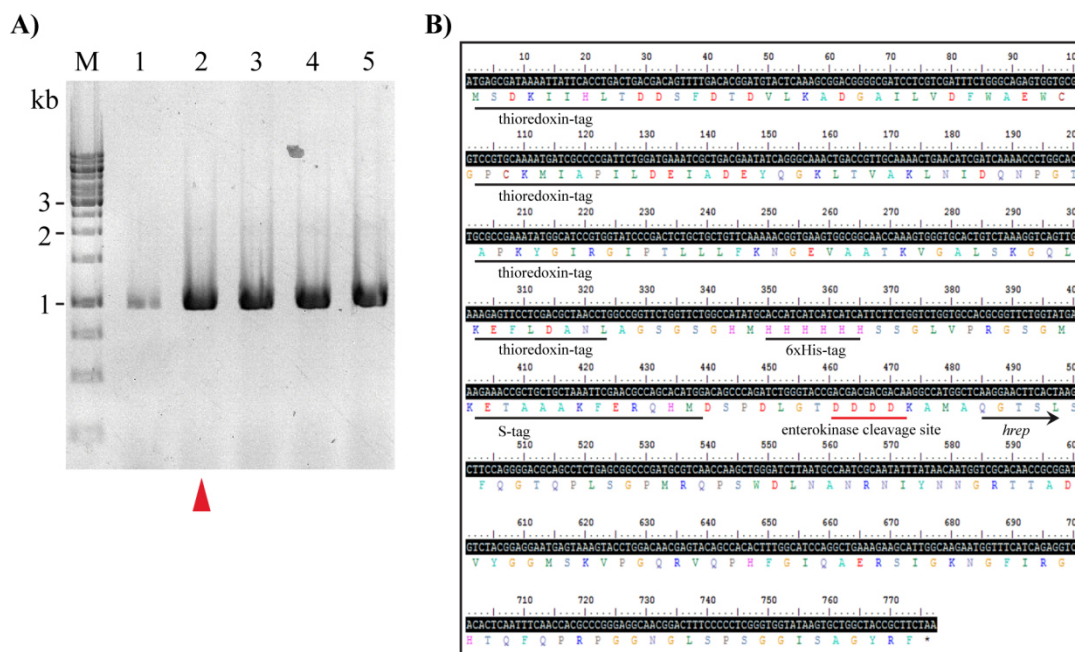
Figure 31 Composition of thioredoxin-S-His6 hymenoptaecin peptide fusion protein. **(A)** PCR amplification of the DNA fragments encoding the multiple cloning site (MCS) of pET32a. The fragments were amplified using the pET T7 oligonucleotides and plasmid DNA from *E. coli* Rosetta 2(DE3) clones transformed with recombinant pET32a was used as a template. The PCR products were separated on a 1.2% agarose gel alongside molecular size markers (lane M, GeneRuler 1 kb DNA Ladder; kb = kilo base pairs) and analysed with HD Green staining (INTAS Science Imaging Instruments GmbH). The major bands (lane 1-5) correspond to the about 1,000 bp MCS containing the sequence coding for the thioredoxin-S-His6 tagged hymenoptaecin peptide. The *E. coli* Rosetta 2(DE3)_pET32a_*hrep* 2 clone was chosen for protein expression and purification (red arrow head). **(B)** The DNA fragment encoding the hymenoptaecin peptide (*hrep*) was incorporated into the multiple cloning site of the plasmid pET-32a and is linked to the thioredoxin-S-His6-tag via the enterokinase cleavage site. The fusion tag consists of a thioredoxin-tag, followed by a 6xHis-tag and an S-tag. The complete sequence encodes a fusion protein thioredoxin-S-His6hld with a molecular weight of about 27.75 kDa. (* = Stop codon)

The expression of the fusion proteins was induced by adding different concentrations of IPTG to the respective cultures of the *E. coli* Rosetta clones, and the expression at 37°C was monitored via SDS-PAGE analyses. The results indicate that the expression of fusion proteins increases over time (Figure 32). For large-scale protein production, *E. coli* cells were harvested 3 hours after induction for the thioredoxin-S-His6hld or 5 hours after induction for the production of the thioredoxin-S-His6hrep when the fusion protein expression was strongest. Before protein purification, solubility of the overexpressed proteins was tested and the fusion proteins thioredoxin-S-His6hld and thioredoxin-S-His6hrep were purified with the BugBuster™ reagent. This reagent is a specifically designed composition of detergents for isolation and separation of soluble and insoluble proteins expressed by bacterial cells. The thioredoxin-S-His6hld and the thioredoxin-S-His6hrep were found in the soluble fraction. Therefore, further protein purification steps could be performed under native conditions.
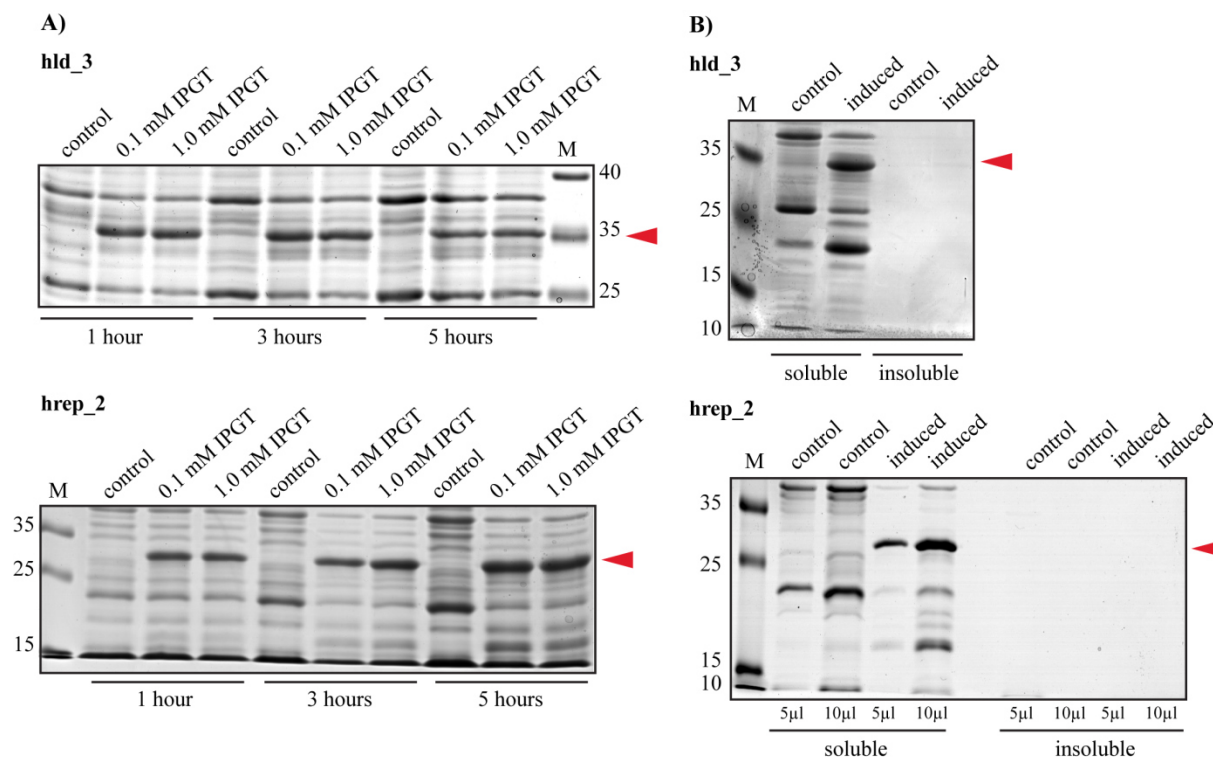
Figure 32 Expression and determination of solubility of thioredoxin-S-His6hld and thioredoxin-S-His6hrep from *E. coli*. **(A)** Expression of hld3 and hrep2 fusion proteins expressed in *E. coli* Rosetta 2(DE3) was induced via addition of different concentrations of IPTG (0.1 mM or 1.0 mM). The bacterial cultures were incubated at 37°C and samples were taken for SDS-PAGE analysis from a non-induced control and the induced cultures after 1, 3 and 5 hours post induction. The heights which correspond to the molecular weights of the fusion proteins are marked with red arrows. **(B)** Extraction of fusion proteins from bacterial cells via BugBuster™. The BugBuster™ reagents were used to isolate hld3 or hrep2 fusion proteins expressed in *E. coli* cells, thereby defining whether the proteins were insoluble or soluble. The thioredoxin-S-His6hld3 as well as the thioredoxin-S-His6hrep2 (indicated by red arrow heads) expressed in *E. coli* are soluble. (M = molecular weight marker; molecular weights in kDa)

4.4.4 Purification of thioredoxin-S-His6 hymenoptaecin fusion proteins via HPLC

Affinity purifications are suitable methods for the isolation of proteins of interest from a protein mixture. Here, *E. coli* cells expressing the fusion proteins were lysed via sonication, and lysates were cleared via centrifugation. Afterwards, cleared lysates were applied to Nickel-NTA columns for His-tag purification in an HPLC based system. Whereas proteins without any affinity to the column resin will pass through the column resin, proteins with a His-Tag specifically bind to the column. In a first step the fusion proteins thioredoxin-S-His6hld and thioredoxin-S-His6hrep were isolated from lysates via His-tag affinity purification (Figure 33). The application of a linear gradient of the elution buffer resulted in the elution of the fusion proteins with increasing imidazole concentrations. Elution fractions containing the thioredoxin-S-His6hld (Figure 33 B) or the thioredoxin-S-His6hrep (Figure 33 C) were taken for further purification in order to remove impurities from the samples.
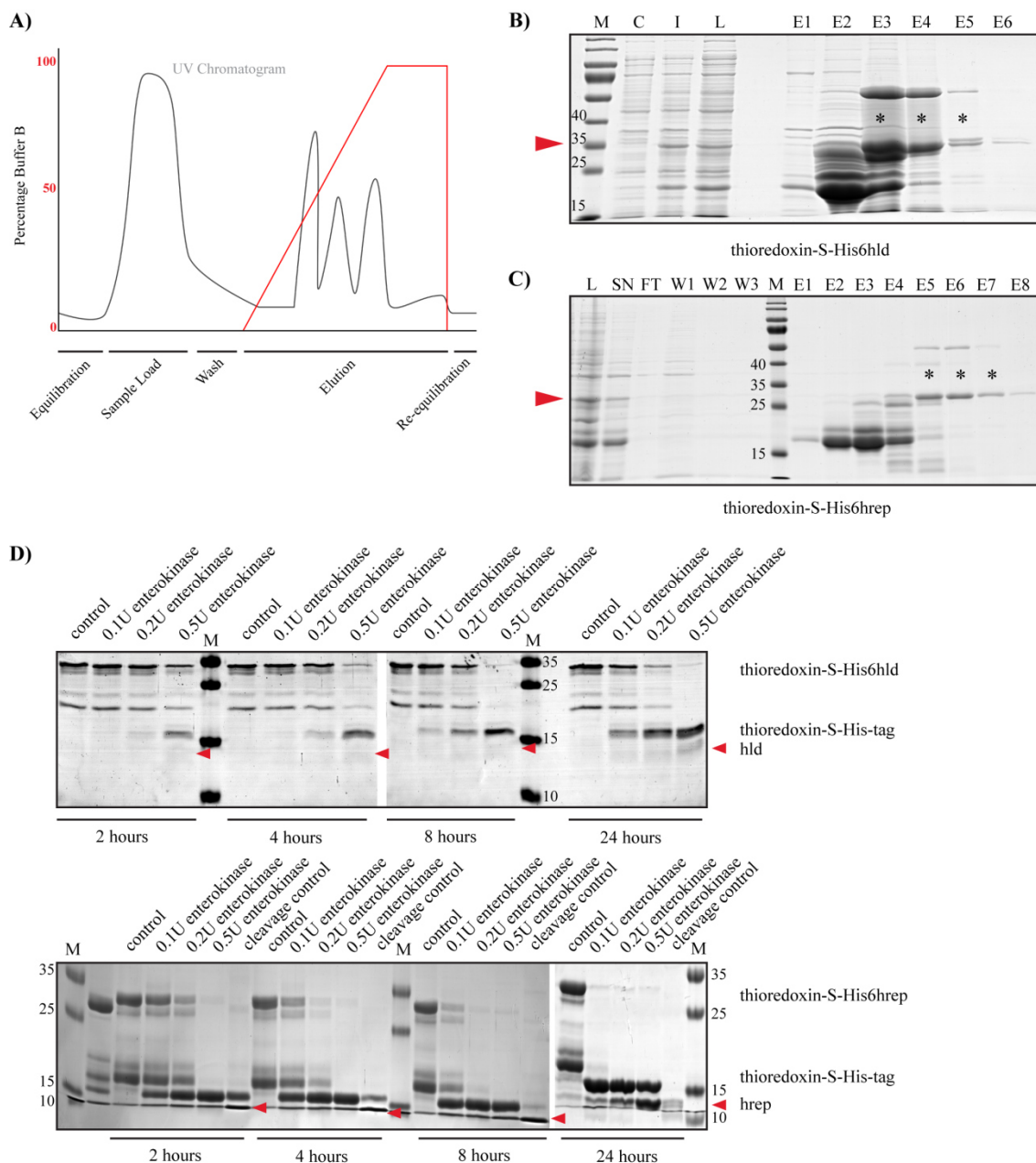
Figure 33 Purification of thioredoxin-S-His6hld and thioredoxin-S-His6hrep from *E. coli* cells. **(A)** The scheme shows the general purification workflow for His-tag purification utilising HPLC (ÄKTA pure). After equilibration of the column material in binding buffer, samples were loaded. The UV chromatogram indicates proteins not bound to the column material, remaining in the flow through. After washing of unspecifically bound proteins, specifically bound proteins are eluted when a gradient of elution buffer with increasing imidazole concentration is applied. Finally, the column material is re-equilibrated. Purification analysis via SDS-PAGE: The **(B)** thioredoxin-S-His6hld fusion proteins and **(C)** thioredoxin-S-His6hrep fusion proteins expressed in 1.0 l culture volume of *E. coli* were purified under native conditions via HPLC (ÄKTA pure). The application of increasing concentrations of imidazole results in the elution of several fractions containing the thioredoxin-S-His6hld or -hrep fusion protein (indicated with a red arrow head). Elution fractions containing the fusion proteins and only few impurities were taken for further purification steps. **(D)** Enterokinase digestion for fusion tag removal. In a small-scale optimisation reaction different amounts of enterokinase were added to 10 µg of the fusion proteins. The digestion was monitored over 24 hours and results were analysed via SDS-PAGE. The complete digestion of the thioredoxin-S-His6 fusion proteins results in the free thioredoxin-S-His6-tag and either the hld peptide or the hrep peptide. Additionally, 10 µg of a cleavage control protein were digested with the enterokinase to confirm functionality of the enzyme. (C = uninduced control culture; I = induced culture; L = lysate; SN = supernatant; FT = flow through; W = wash fraction; E = elution fraction; M = molecular weight marker; molecular weights in kDa; * = elution fractions chosen for further purification)

Assuming that the affinity-tag would influence functionality and specificity of the hymenoptaecin peptides, the tags were removed with the enzyme enterokinase. In a small-scale optimisation reaction different amounts of enterokinase were added to each 10 μg of the fusion proteins and samples were taken after 2, 4, 8 and 24 hours (Figure 33). The complete digestion of the thioredoxin-S-His6 fusion proteins results in the release of the thioredoxin-S-His6-tag and either the hld peptide (about 12.3 kDa) or the hrep peptide (about 10.7 kDa) (Figure 33 D). For large-scale digestion ratios of 50 Units enterokinase per mg thioredoxin-S-His6hld and 20 Units enterokinase per mg thioredoxin-S-His6hrep proved to be efficient. Whereas the hld fusion protein was digested for 24 h, digestion of the hrep fusion proteins over 8 hours already resulted in sufficient cleavage, and results could not be improved by increased digestion time. Separation of the remaining fusion-tags, uncleaved fusion proteins and of other impurities from both hymenoptaecin peptide samples (hld and hrep) was achieved with the use of cation exchange chromatography. A protein's isoelectric point (pI) is the pH value at which the protein has no net charge. At a buffer pH below the protein's pI, the protein will have a net positive charge. The ion exchange resins for cation exchange chromatography are negatively charged with an affinity for molecules having net positive surface charges. For cation exchange chromatography, samples containing the cleaved hymenoptaecin peptides were dialysed to adjust the pH value to a neutral level (6.8). Both hymenoptaecin derived peptides exhibit a high pI of 10.62 (hld) and 11.27 (hrep). Therefore, it was assumed that the net positively charged peptides specifically bind to the negatively charged column resin. The samples were loaded on the HisTrap CaptoS columns and the ion exchange was performed. The results show that major impurities of higher molecular weight around 15 kDa could be removed (Figure 34 A and B), and the eluates containing the proteins of interest were concentrated. The SDS-PAGE analyses also indicate that the hld sample contained an additional protein with a molecular weight of about 40 kDa, and two proteins with molecular weights around 10 to 12 kDa (Figure 34 C). The predicted size of the recombinant hld is about 12.3 kDa. Further, the hrep sample contained small amounts of proteins with a molecular weight around 13 to 15 kDa and a protein only slightly smaller than 10 kDa. Several protein bands were analysed via mass spectrometry. However, the identification of the purified proteins via mass spectrometry was inconclusive. The tryptic digestion of the proteins might have resulted in less than one peptide per protein due to their low molecular weight. Therefore, the samples did not meet an essential condition for the identification of proteins via MS/MS mass spectrometry since at least two peptides are necessary for a significant identification.

The concentrations of purified recombinant hymenoptaecin peptides hld and hrep expressed in *E. coli* were determined via the measuring tool of ImageJ. In spite of fusion proteins being isolated and purified out of several litres of bacterial culture (5 litres each), only small amounts of the peptides were obtained. The first limiting factor appears to be the availability and efficiency of the enzyme enterokinase for the digestion of fusion proteins. Overall, 4 mg thioredoxin-S-His6hld and 10 mg thioredoxin-S-His6hrep fusion protein were digested. After the subsequent purification steps,

hymenoptaecin peptide samples with concentrations of only 10.7 μM (hld) and 111.36 μM (hrep) were obtained suggesting severe loss of peptides during the digestion reaction and following purification steps. In spite of higher peptide stock solution concentrations that are generally required for functional assays (Casteels et al., 1993; Casteels et al., 1990; Tonk et al., 2014; Tonk et al., 2015a; Tonk et al., 2015b), both samples were kept under sterile conditions and used in growth inhibition assays with two different bacterial test strains.
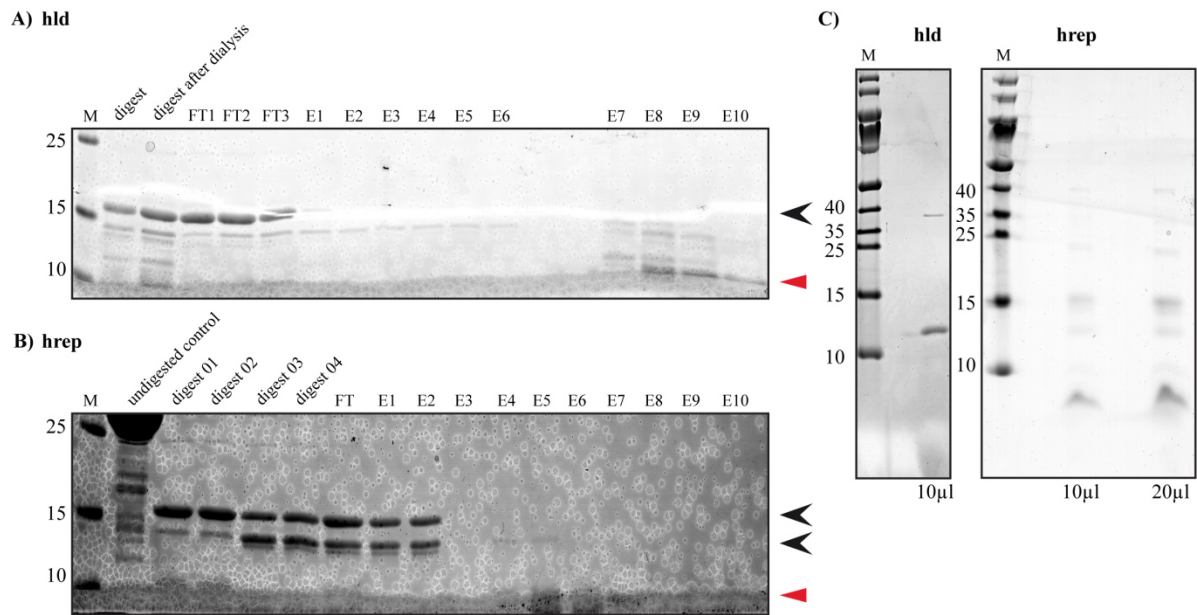
Figure 34 Cation exchange for further purification of protein samples. The sample containing A) the thioredoxin-S-His6-tags and hld or B) the thioredoxin-S-His6-tags and hrep after digestion of the fusion proteins were loaded on a cation exchanger (HisTrap CaptoS) for separation of the hymenoptaecin peptides from the tags. The peptides exhibit a net positive charge within a buffer with neutral pH and therefore should specifically bind to the negatively charged column resin. Increasing the salt concentration within the buffer leads to elution of proteins and peptides bound to the resin. Peptides considered to be hymenoptaecins hld (E8 - E10) or hrep (E5 and E6) due to the predicted molecular weights are marked with red arrow heads. Additional proteins within each sample are marked with black arrow heads. C) After the ion exchange elution fractions containing small peptides with molecular weights around 10-12 kDa were concentrated and checked for impurities via Tricine-SDS-PAGE. (FT = flow through; E = elution fraction; M = molecular weight marker; molecular weights in kDa)

4.4.5   Functional analysis of hymenoptaecin peptides

The functional characterisation of the purified recombinant *C. floridanus* hymenoptaecin peptides was performed following a previously published method (Tonk et al., 2014; Tonk et al., 2015b). Briefly, overnight cultures of two bacterial test strains, *E. coli* D31 and *S. aureus,* were diluted to an OD_{600} of 0.001 and transferred to a 384 well plate. The dilution of the bacterial cultures to a low OD_{600} guaranteed full contact of the bacteria with the peptides at the beginning of the measurement. The peptide solutions were added per well in decreasing concentrations given in Table 17. The well plates were incubated at 37 °C for 20 hours and OD_{600} was measured every 15 min (900 sec). The resulting growth curves of the respective bacteria incubated with or without the peptides were obtained. One

replicate experiment for each bacterial strain was performed and the results are described in the following.

Table 17 Concentrations of hymenoptaecin peptides for antimicrobial assays. Samples of purified hymenoptaecin peptides hld or hrep as well as a 1:1 mixture of both peptides were applied to bacterial cultures and bacterial growth was analysed. Therefore stock solutions were diluted and concentrations [µM] of each peptide for the tested samples (A-I) are listed below.

				Peptide concentration [µM]					
	A	B	C	D	E	F	G	H	I
hld	2.68	1.34	0.67	0.33	0.18	0.08	0.04	0.02	0.01
hrep	27.8	13.9	6.96	3.48	1.74	0.87	0.44	0.22	0.11
hld + hrep	1.34	0.67	0.33	0.18	0.08	0.04	0.02	0.01	0.005
	13.9	6.96	3.48	1.74	0.87	0.44	0.22	0.11	0.05

Each bacterial growth curve is characterised by four main phases: the initial lag phase, when the freshly inoculated bacteria adapt to the new environment and do not replicate; the logarithmic (log) phase, when bacteria replicate and divide in a logarithmic manner; the stationary phase, when a netto growth rate of zero can be observed since the number of reproducing cells is equal to the rate of cell death; and the death phase, when no reproduction but cell death occurs and the number of living cells constantly decreases.

Most growth curves of control and treated cultures of *E. coli* D31 exhibited a lag phase of about 4 hours (about 14,000 sec) (Figure 35). In the lag phase the bacterial growth stagnated completely and even a decline in the number of living cells was observed. After two to three hours a minimum OD_{600} was measured followed by the recovery of bacterial growth. This phenomenon was more distinct in cultures *E. coli* D31 treated with the hymenoptaecin derived peptides, especially the hrep and the mixture of both peptides, which may indicate a hymenoptaecin peptide mediated reduction of the number of living *E. coli* D31 at the beginning of the measurement. Due to the low number of bacteria the full contact of the cells with the peptides is very likely and bacteria might be exposed to the putative antimicrobial peptides especially at the beginning of each experiment. After about four hours, the bacteria entered the log phase lasting about 7 h (about 25,000 sec) (Figure 35). Almost all cultures entered the stationary phase after about 11 h (about 40,000 sec) (Figure 35).

The inhibition of growth appeared to be depending on peptide concentrations. Whereas better growth in comparison to untreated controls was observed in samples treated with higher concentrations of hrep (> 3.48 µM; Figure 35 A-C), the application of hrep peptide concentrations of 0.87-3.48 µM (Figure 35 D-F) resulted in growth similar to the untreated control. A massive elongation of the lag phase was observed in the *E. coli* D31 sample treated with 0.33 µM hld. The bacteria entered the log phase only after 11 h and did not enter the stationary phase within the duration of the experiment (Figure 35 D). In samples treated with 0.67-1.34 µM hld a complete growth inhibition was observed (Figure 35 B and C). However, highest concentrations of hld (2.68 µM) did not inhibit bacterial

growth, but treated bacteria reached a higher plateau in the stationary phase. A mixture of both hymenoptaecin derived peptides seemed to exhibit a mild growth inhibition on *E. coli* D31 at hrep concentrations of 0.87-6.96 µM and hld concentrations of 0.08-0.67 µM (Figure 35 B-E). Even though the cultures recovered from the growth inhibition in the lag phase and the further growth curves showed a normal course the overall OD_{600} after 20 h was distinctly lower than the OD_{600} of the untreated controls. It can be assumed that hld exhibits bacteriostatic effects on *E. coli* D31, as it has been described for example for several defensins from the tick *Ixodes ricinus*. Defensin peptides causing a growth plateau (stationary phase) lower than that of the negative control were considered to have bacteriostatic effects (Tonk et al., 2015a). Concentrations < 0.08 µM for hld or < 0.87 µM for hrep did not influence bacterial growth (data not shown).
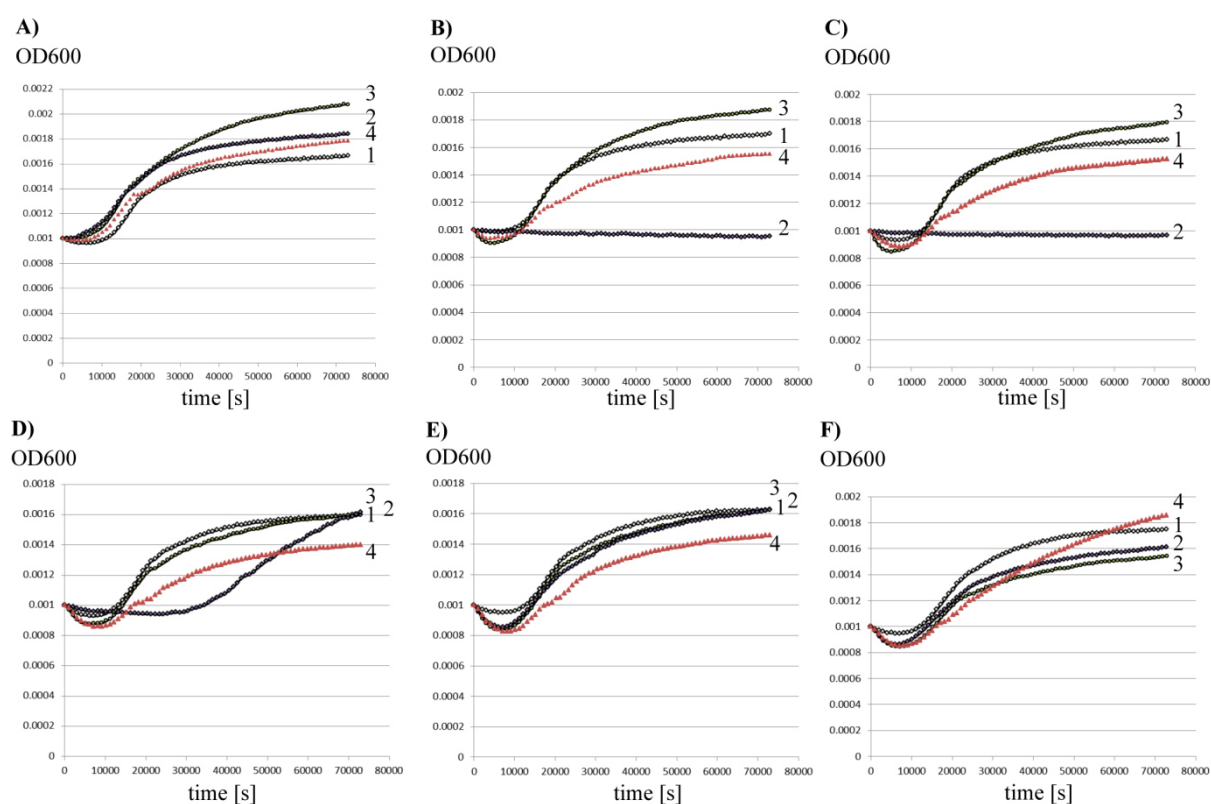
Figure 35 The functional analysis of hymenoptaecin peptides. The growth of *E. coli* D31 was monitored in the absence or presence of recombinant hymenoptaecin peptides. An overnight culture of *E. coli* D31 was diluted to a starting optical density OD_{600} of 0.001 and either peptide sample buffer (1; untreated control) or different concentrations (decreasing from **(A) - (F)**) of either hld (2), or hrep (3) or a 1:1 mixture of both peptides (4) were added. Samples were incubated at 37°C with constant shaking and OD_{600} was measured every 900 sec (15 min) over a period of 20 hours. All peptide concentrations are given in (Table 17).

The recombinant hld and the mixture of both hymenoptaecin derived peptides exhibit inhibitory effects on the growth of *S. aureus* similar to the effects on *E. coli* D31 described above. Again, a stronger reduction of the optical density during the lag phase was observed in treated cultures when compared with untreated controls (Figure 36). Furthermore, distinct bacteriostatic effects, with *S. aureus* reaching lower growth plateaus in treated cultures than in untreated control cultures, occurred when *S. aureus* was treated with hld peptide concentrations of 0.01-2.68 µM, or a mixture of both peptides with hrep concentrations ≥ 0.05 µM and hld concentrations ≥ 0.005 µM (Figure 36 A-I).

However, the bacteriostatic effect of the peptide mixture on *S. aureus* might be more a result of the hld peptide. Treatment with the recombinant hrep peptide alone did not result in an inhibition of bacterial growth (Figure 36 A-I).
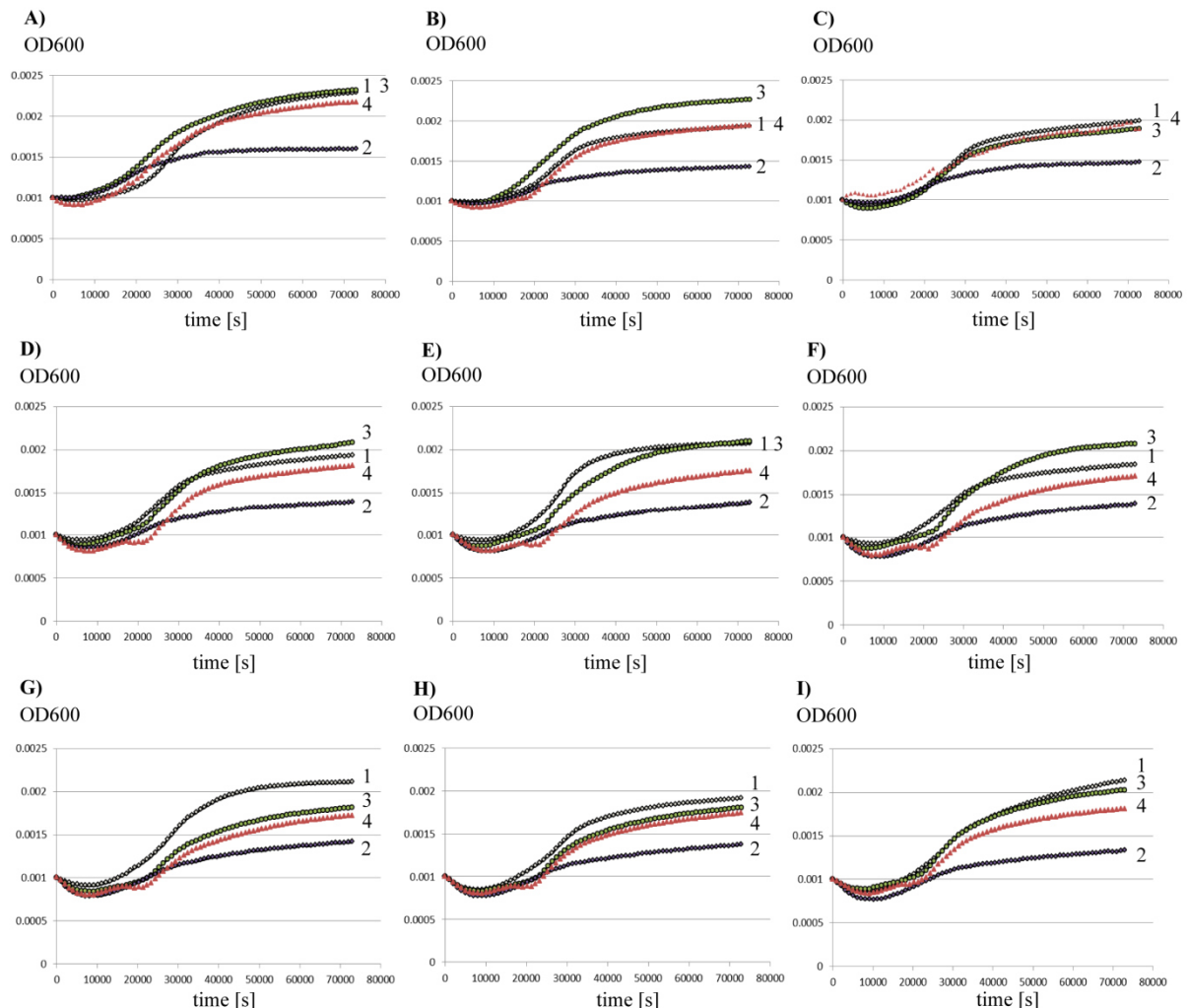

Figure 36 The functional analysis of hymenoptaecin peptides. The growth of *S. aureus* was monitored in the absence or presence of recombinant hymenoptaecin peptides. An overnight culture of *S. aureus* was diluted to a starting optical density OD_{600} of 0.001 and either peptide sample buffer (1; untreated control) or different concentrations (decreasing from **(A) - (I)**) of either hld (2), or hrep (3) or a 1:1 mixture of both peptides (4) were added. Samples were incubated at 37°C with constant shaking and OD_{600} was measured every 900 sec (15 min) over a period of 20 hours. All peptide concentrations are given in (Table 17).

In a previous study a hymenoptaecin peptide (hrep) was purified under denaturing conditions and tested against different bacterial strains. The peptide showed activity only against *E. coli* D31 when tested in an agar diffusion assay (Kupper, 2011). In this work, the hymenoptaecin derived peptides hld and hrep were purified under native conditions assuming that avoiding extensive refolding procedures results in peptides with close to native functions. Overall, these first functional analyses showed bacteriostatic effects of the purified recombinant hymenoptaecin peptide hld. The results suggest that the recombinant hld and hrep exhibit different specificities against the tested bacteria. Whereas hld and a mixture of both peptides may have bacteriostatic effects on *E. coli* D31 and hld showed bacteriostatic effects against *S. aureus*, treatment with the hrep peptide did not lead to a significant inhibition of *E. coli* D31 or *S. aureus* growth.

4.5   The distribution of *B. floridanus* in ovarian tissue of *C. floridanus* during host oogenesis

4.5.1   The tissue localisation of *Blochmannia* within an ovariole of *C. floridanus*

The endosymbiosis between the ant *C. floridanus* and the bacterial endosymbiont *B. floridanus* exhibits characteristic features which are results of the coevolution of host and symbiont over approximately 40 million years (Degnan et al., 2004; Sameshima et al., 1999; Sauer et al., 2000). These features display not only the complex metabolic interactions between the ant and the bacteria or the already extensively investigated highly dynamic distribution of *B. floridanus* in the midgut tissue during the ant host's metamorphosis (Feldhaar et al., 2007; Stoll et al., 2010). Furthermore, as the complete development of the ants depends on the symbiosis (Zientz et al., 2006), the vertical transmission of the endosymbionts is of inevitable importance. The maternal transmission secures the transfer of *Blochmannia* to the next generation. However, the knowledge about the vertical transmission in *C. floridanus* remained limited until last. In the following, the distribution of *B. floridanus* in the ovarian tissue of *C. floridanus* during the ant's oogenesis will be described for the first time.
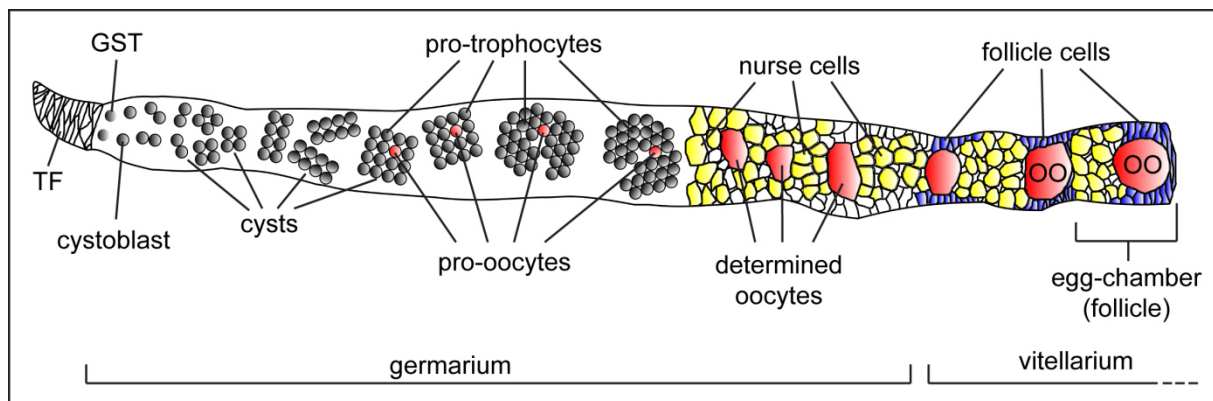

Figure 37 Schematic diagram of oogenesis in a *C. floridanus* ovariole. The three main parts of the ovariole are the terminal filament (TF), which positions the ovariole within the ant's abdomen, the germarium and the vitellarium. The apical germarium following the TF harbours germ-line stem cells (GSTs) and their immediate progeny, the cystoblasts. The latter undergo a series of mitotic divisions and differentiate into cystocytes. The cystocytes form cysts and remain connected via ring canals. Within the 16-cell and 32-cell cysts, beginning differentiation of cystocytes allows the discrimination of pro-oocytes and pro-trophocytes. Whereas one pro-oocyte of each cyst will develop into the determined oocyte and later on oocyte (OO), pro-trophocytes develop into nurse cells. In the basal germarium, nurse cell packages form functional clusters with oocytes which become increasingly surrounded by follicle cells. These so-called egg-chambers (follicles) enter the vitellarium, where the oocytes undergo growth and vitellogenesis and finally enter maturation by reinitialisation of meiosis.

The pairwise ovaries of *C. floridanus* exhibit two bundles of ovarioles in varying numbers depending on the caste. The number of ovarioles observed in *C. floridanus* workers ranged from 2 to 12, and they are of the polytrophic-meroistic type. The three main parts of each ovariole are the terminal filament, the germarium and the vitellarium (Figure 4 and Figure 37). The apical part of each germarium harbours the germ-line stem cells, which are located close to the terminal filament, and their direct progeny, the so-called cystoblasts. Within the germarium, the progeny of germ-line stem cells undergo a series of synchronised mitotic divisions and differentiate into clusters of cystocytes, the cysts. In the

lower or basal part of the germarium one cystocyte of a 32-cell cyst enters meiosis and differentiates into the oocyte, whereas the remaining 31 cystocytes develop into nurse cells (trophocytes) as it has been described for several hymenopterans including ants such as *Aphaenogaster rudis*, *Myrmica americana*, and *Lasius niger* (Billen, 1985; Khila and Abouheif, 2008; King and Büning, 1985; Klag and Bilinski, 1994; Tanaka and Hartfelder, 2004). Egg-chambers (follicles) are clusters of one oocyte and 31 supporting nurse cells and enter the vitellarium, where the growing oocyte undergoes vitellogenesis. Older follicles are shifted towards the posterior end of the ovariole due to the formation of younger egg-chambers.

In this work, Alexa488 fluorophore labelled 16S rRNA specific probes for *B. floridanus* (Bfl172_Alexa488) were used to identify the endosymbionts within the ovarioles of queenless *C. floridanus* workers. Notably, the abundance and distribution of the endosymbiont within the ovariolar tissue appeared quite variable with respect to the events of oogenesis.

The terminal filament is the anterior anchor of each ovariole in *C. floridanus* and positions the ovariole inside the ant's abdomen together with the ovariolar epithelial sheath and the peritoneal sheath. The terminal filament is followed by the apical germarium (Figure 38) which harbours the germ-line stem cells (GSCs). Following differential mitosis and division of one GSC, two cells are produced. While one cell retains the totipotent character of the parent cell and remains a GSC, the other cell, the cystoblast, undergoes mitotic cycles and divides into cystocytes. Within the germarium, germ cells and their progeny are accompanied by numerous somatic cells, as it has been described for other ants (Billen, 1985; Khila and Abouheif, 2008). Interestingly, this most apical part of the germarium appeared to be an ovariolar segment free from *Blochmannia* (Figure 38), as Bfl172_Alexa488-positive green fluorescence could not be detected. These observations were confirmed by a 3-dimensional reconstruction of a stack of pictures taken in the z-axis of a germarium. The generated Leica Image File was rendered into a 3-D model using the Volume viewer of ImageJ 1.49k, which allowed the analysis of the tissue's full volume (Figure 38 D) showing the absence of *Blochmannia*-positive fluorescence in the apical germarium.
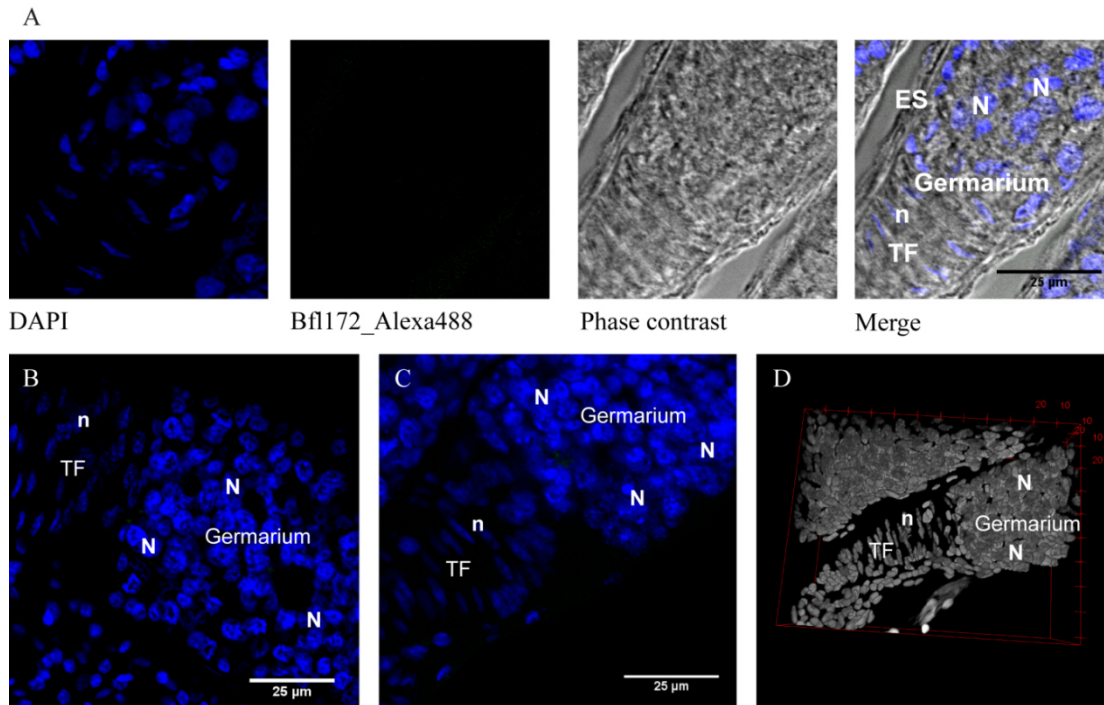
Figure 38 The apical germarium within the ovarioles of *C. floridanus*. The germarium follows the terminal filament (TF) at the anterior tip of each ovariole. Its apical part contains germ-line stem cells and their direct progeny, the cystoblasts, which are accompanied by numerous somatic cells. Each ovariole is surrounded by the ovariolar epithelial sheath (ES). **(A) - (D)** Several tissue samples of *C. floridanus* ovaries were analysed via confocal laser scanning microscopy. Each apical part of the germarium appeared to be *B. floridanus* free since the Bfl172_Alexa488 positive fluorescence (green) was absent. **(A)** A series of pictures taken of the same position of an apical germarium shows the fluorescence images of the channels DAPI (blue) and Bfl172_Alexa488 (green) and the phase contrast image. The overlay (Merge) of the pictures represents the whole structure of the apical germarium. **(B)** and **(C)** The pictures represent merged fluorescence images (DAPI in blue; Bfl172_Alexa488 in green) of the apical germarium of different ovarioles. **(D)** A 3D reconstruction of a z-stack of pictures of a germarium allowed the analysis of the tissue's full volume confirming the absence of *B. floridanus*. DAPI (blue in (A) - (C); grey in (D)) was used for counterstaining of DNA (n = nuclei of TF cells; N = nuclei of cells of the germarium).

Within the germarium, the cystocytes undergo several synchronised mitotic divisions thereby producing 2-cell, 4-cell, 8-cell, 16-cell and finally 32-cell cysts. While the younger and more apically cysts are formed, older cysts are shifted towards the lower part of the germarium, following the anterior-posterior body axis of the ants. In the segment containing cysts after the 4th and 5th division of cystocytes, Bfl172-positive fluorescence and therefore a quite sudden infection of the tissue with the endosymbionts was detected (Figure 39 A). Interestingly, the infection appeared to start in the outer layers of the ovariolar tube. Only cells lying under the ovariolar epithelial sheath seemed infected, whereas the inner layers of the ovariole harbouring the cystocytes did not harbour *B. floridanus* (Figure 39 B). However, the infection proceeded rapidly to the centre of the ovarioles. A stack of pictures was captured in the z-axis of an ovariole sample and rendered as Leica Image File format for processing. The stack was then opened on Fiji as non-flattened files and channels were separated. The background was subtracted using the rolling ball background subtraction, and the different channel images were merged to a stacked file. The file was then radially re-sliced in a "top to bottom" fashion to change the perspective of view. The re-sliced file consisting of 2,804 slices was then rendered into a 3-D model in the chosen perspective using the Volume viewer function of Fiji. To generate the sliced

model, 252 blank frames were inserted at regular intervals, and the model was restacked in the chosen perspective allowing the view into the tissue. The endosymbionts appeared to spread into the whole tissue within a region of approximately 100 µm width following the apical-basal axis of the ovariole (Figure 39 C). Soon the massive infection of the whole tissue was observed, albeit cystocytes of the 16-cell cysts within the anterior "infection zone" and 32-cell cysts within the more posterior "infection zone" remained uninfected at first (Figure 39 D). These cystocytes could be easily distinguished from cells of somatic origin. Whereas the latter were quite small, the round cystocytes exhibited a characteristically expanded cytoplasm (Figure 39 B and D), within which *B. floridanus* was not detected.
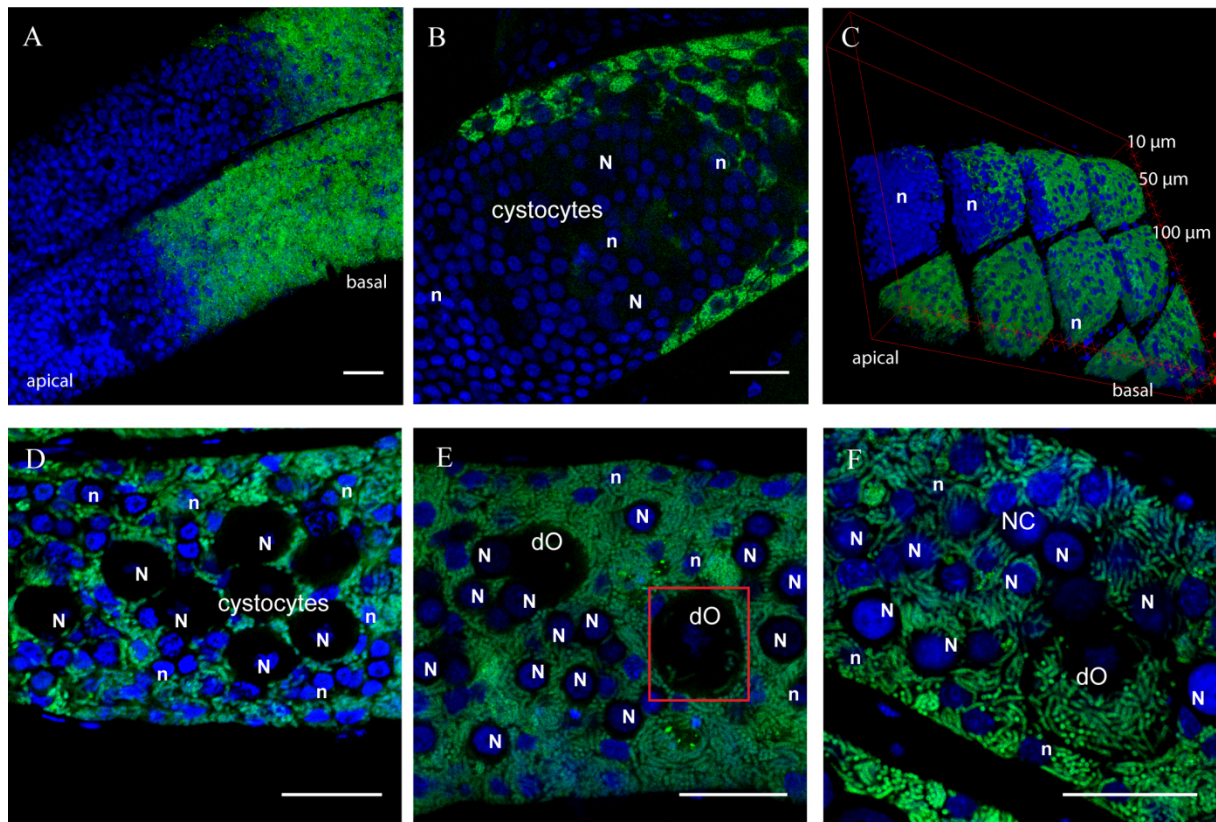

Figure 39 The infection of the basal germarium of *C. floridanus* ovarioles with *B. floridanus*. Green colour indicates Bfl172_Alexa488 positive signals and therefore *B. floridanus* detection, and blue colour marks nuclei stained with DAPI. All pictures show segments of the lower germarium and are orientated from left to right following the anterior-posterior body axis. **(A)** The overview shows the sudden appearance of *B. floridanus* in the basal part of the germarium, whereas the more apical segment remained bacteria-free. **(B)** At the beginning of the "infection zone" only the outer layers of the ovariole were infected, whereas the inner layers where the cystocytes reside were free from *Blochmannia*. **(C)** The 3-D model of ovariole segments shows the progressing bacterial infection of the whole ovariole tissue. A stack of pictures was taken in the z-axis of the sample and radially re-sliced in a "top to bottom" fashion using Fiji. The resulting stack was rendered into a three-dimensional model using the Volume viewer function of Fiji. 252 blank frames were inserted at regular intervals to create the sliced model. The side view into several positions of the tissue shows the spreading of *B. floridanus* from the outer layers towards the inner layers of the ovariole. **(D)** Even though *B. floridanus* appeared to rapidly infect the whole tissue, the cytoplasm of cystocytes forming 16-cell cysts and 32-cell cysts remained bacteria-free. **(E)** Only after the 5th division, one of the cystocytes differentiates into the determined oocyte (dO). At this stage scattered endosymbionts were observed in the cytoplasm of the dOs (marked with red square). **(F)** Subsequently, the dOs appeared to become increasingly infected with *B. floridanus*, while egg-chambers consisting of nurse cells (NC) and the dO, and later oocyte, are formed. Scale bars represent 25 µm. (N = nuclei of nurse cells; n = nuclei of cells of somatic origin)

Following the 4th mitotic division, differentiation of the cystocytes of a 16-cell cyst is initiated. Several cystocytes of a 16-cell cyst may enter meiosis and are termed pro-oocytes due to their potential of becoming oocytes. Nevertheless, within the 32-cell cysts, only one pro-oocyte is selected to differentiate into the determined oocyte (dO), as it has been described for the oogenesis in polytrophic-meroistic ovarioles in *Drosophila* or the hymenopteran *Coleocentrotus soldanskii* (Gonzalez-Reyes et al., 1997; Grell and Generoso, 1982; Klag and Bilinski, 1994). The dO exhibits a characteristic phenotype as it grows faster than the other 31 cystocytes which become pro-trophocytes and eventually develop into nurse cells (trophocytes). Accompanied by somatic follicle cells, the dO is shifted towards the centre of the ovariolar tube and reaches a size of at least six nurse cells. At this stage, scattered *B. floridanus* were detected inside the dO's cytoplasm for the first time (Figure 39 E). Therefore, it can be stated that the *C. floridanus* oocytes are initially infected with the bacterial endosymbionts very early in oocyte development similarly to the oocyte infection as described for *C. ligniperda* (Blochmann, 1882; Buchner, 1918). In *C. floridanus*, the initial infection of the oocytes was observed within the basal germarium of each ovariole. Subsequently, determined oocytes became increasingly infected with *B. floridanus* during egg-chamber formation (Figure 39 F). These functional clusters consisting of nurse cells (NCs) followed by the oocyte, which is surrounded by follicle cells, eventually enter the vitellarium (Figure 40).

During growth of the oocyte, each egg-chamber is shifted towards the posterior end of the ovariole when younger egg-chambers are formed. Hence, the vitellarium appears as a linear arrangement of egg-chambers. At the beginning of the vitellarium the whole tissue appeared to be infected with *B. floridanus*, as Bfl172_Aleax488 signals were found in the oocyte but also surrounding nuclei of NCs and follicle cells (Figure 40 A). However, under the experimental conditions for FISH and confocal laser scanning microscopy in this work, it was not possible to define the exact cellular localisation of the bacterial endosymbionts. Nevertheless, the number of *Blochmannia* within the oocyte and the surrounding follicle cells appeared to change in a reversed manner. While the ooplasm became increasingly packed with the endosymbionts, the number of *B. floridanus* decreased continuously in the follicle cells. Eventually, follicle cells were free from *Blochmannia,* whereas the endosymbionts occupied the whole ooplasm of younger oocytes (Figure 40 B-D). In this segment of the vitellarium, the periodical alternation of bacteria-filled oocytes and bacteria-free NC packages resulted in a necklace-like appearance of the tissue (Figure 40 G). During the continuing growth of the oocyte, follicle cells are involved in the formation of the egg shell at the end of the vitellogenic growth phase (Büning, 2005). In this stage of vitellogenesis, when the oocytes grow massively, the bacterial endosymbionts seemed to be shifted towards the posterior pole of the oocyte (Figure 40 D-F). In fully grown oocytes, large numbers of *B. floridanus* were detected only at their posterior pole (Figure 40 F).
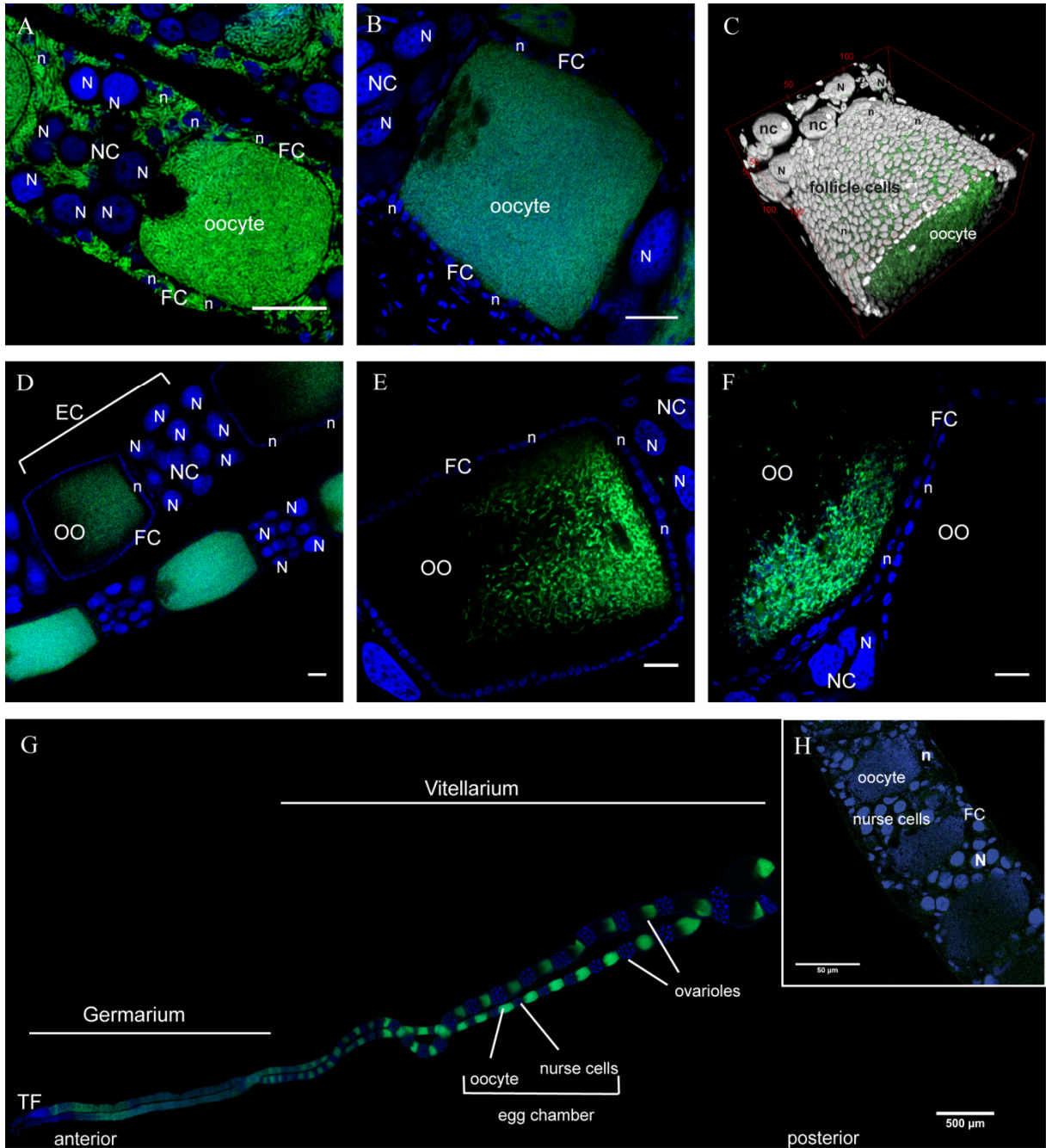
Figure 40 The distribution of *B. floridanus* within the vitellarium of *C. floridanus* ovarioles. Egg-chambers (EC) consisting of a nurse cell (NC) package followed by one oocyte (OO), which is surrounded by a single layer of follicle cells (FC), are formed in the germarium and shifted towards the vitellarium when younger ECs are formed. **(A)** At the stage of early egg-chamber formation, the whole tissue including oocytes and follicle cells appeared to be infected with *B. floridanus*. **(B)** and **(C)** During oocyte growth the ooplasm appeared to become increasingly packed with the endosymbionts, whereas *B. floridanus* was no longer detectable in follicle cells. The snapshot of the 3-D model of a young oocyte shows that *B. floridanus* occupies the whole ooplasm. **(D) - (F)** However, during the growth of the oocytes the bacterial endosymbionts were localised continuously towards the oocyte's posterior pole. **(F)** In fully grown oocytes, *B. floridanus* resided within a small volume at the posterior pole. **(G)** The overview presenting two ovarioles summarises the distribution of *B. floridanus* during host oogenesis. Whereas the terminal filament and the apical germarium appeared to be not infected, *B. floridanus* seems to rapidly infect the basal germarium. In the vitellarium, the amount of endosymbionts in the OOs increases continuously. The periodical alternation of bacteria filled oocytes and bacteria free NC packages results in a necklace-like appearance of the ovarioles. **(H)** A representative view of a section of an ovariole treated with the negative control probe Bflo172sense (Alexa488, green) and DAPI (blue). The sense probe did not detect bacteria residing within the oocytes. Scale bars in (A) - (F) represent 25 µm; 500 µm in (G); and 50 µm in (H). (N = nuclei of nurse cells; n = nuclei of cells of somatic origin and follicle cells; Bfl172_Alexa488 = green; DAPI = blue or grey).

As mentioned above, the detection of *B. floridanus* with fluorescent RNA specific probes using confocal laser scanning microscopy for visualisation could only provide information about the general distribution of the bacterial endosymbiont within the tissue. However, the exact intracellular or possible extracellular localisation of *B. floridanus* was not resolved. Hence, the results of a detailed qualitative analysis of ovariole tissue sections utilising transmission electron microscopy will be described in the next chapter.

4.5.2 A detailed analysis of the cellular localisation of *Blochmannia* in ovarian tissue of *C. floridanus* using transmission electron microscopy

Previously published results of a TEM analysis of the localisation of *B. floridanus* in *C. floridanus* ovarian tissue could only show the massive abundance of the bacterial endosymbionts within the ooplasm of fully grown oocytes shortly before oviposition, whereas the follicle cells and nurse cells are bacteria free (Sauer et al., 2002). In this work, entire ovarioles were successfully dissected and treated following the classical fixation protocol for transmission electron microscopy which is especially suitable for membrane structure preservation. In the following, the intracellular localisation of the endosymbionts throughout several stages of oogenesis will be described in detail.

The detection of fluorescent dye labelled probes using confocal laser scanning microscopy is technically limited, as a minimal fluorescence signal is necessary for positive detection. Hence, the analysis of tissue sections via TEM appeared to be a suitable method to confirm the absence of *B. floridanus* in the apical germarium of *C. floridanus* (Figure 38). Several longitudinal sections of different ovariole samples were analysed focussing on the identification of the rod shaped bacteria with a width of about 1 μm and varying length. Neither the cells of the terminal filament nor the germ-line stem cells (GSCs) nor their direct progeny, the cystoblasts (CB), were infected with *B. floridanus* (Figure 41). The GSCs and CBs exhibit characteristically large nuclei, and the cytoplasm which contains noticeable mitochondria appears electron-dense and therefore quite dark (Figure 41 C and E). However, the endosymbionts were never found in the cytoplasm of these germ cells. Additionally, small somatic cells lying directly under the basal lamina and ovariolar epithelial sheath were not infected with *B. floridanus* (Figure 41 D). Several holes disrupting the integrity of the tissue sections of the germarium are marked with squares. However, these artefacts did not influence the display of the described structures.

As described above, cystoblasts undergo several synchronised divisions, and the resultant cystocytes remain attached to each other by intercellular bridges forming clusters, the so-called cysts. The continuous formation of new cysts leads to the progressing shift of older cysts towards the lower or basal segments of the germarium. Eventually, 16-cell and 32-cell cysts are located in the segment described as the "infection zone" (chapter 4.5.1). In this segment, the bacterial endosymbionts were observed in the cytoplasm of cells of the ovariole's outer layer, where they initially infect the tissue, which confirms the results obtained via FISH.
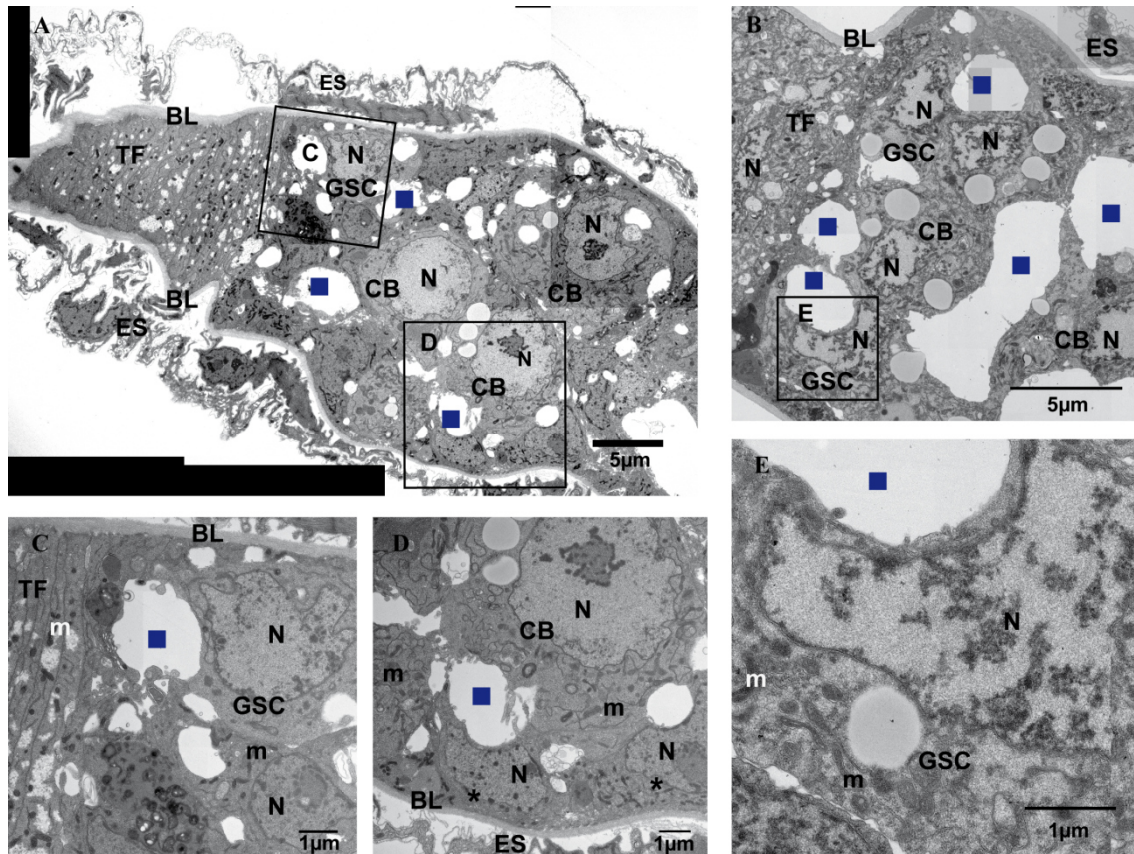
Figure 41 The ultrastructure of the apical germarium in *C. floridanus* ovarioles. **(A)** and **(B)** The electron micrographs of two longitudinal sections of the anterior tips of two ovarioles give an overview of the terminal filament (TF) and the apical germarium. The germ-line stem cells (GSCs) are in close contact with the TF. The basal lamina (BL) and the ovariolar epithelial sheath (ES) surround each ovariole. **(C)** and **(E)** Each enlargement of either (A) or (B) shows a GSC exhibiting its large nucleus (N). The cytoplasm of both GSCs contains numerous mitochondria (m) but appears to be free from *B. floridanus*. **(D)** Enlargement of a cystoblast (CB), the direct progeny of GSCs. The CB is in close contact to small somatic cells lying directly under the BL and ES (marked with asterisks). Throughout the apical germarium no *B. floridanus* infection was observed. The cytoplasm of GSCs, of CBs and of small somatic cells appeared bacteria-free. Holes in the tissue sections which are likely to be artefacts of fixation and Epon embedding are marked with blue squares.

Additionally, the detailed analysis with TEM revealed that these small somatic cells, lying directly under the basal lamina, harboured *B. floridanus* which were located exclusively intracellularly either free in the cytoplasm or in vacuole-like structures (Figure 42 A-C and F). Here, the rod-shaped bacteria were easily distinguishable due to their spotted appearance and the characteristic Gram-negative cell wall structure, as the two bacterial membranes were clearly visible (Figure 42 B and C). In this segment of the early "infection zone", cystocytes could be easily discriminated from somatic cells. Whereas the cytoplasm of the latter appeared comparatively bright, the cytoplasm of cystocytes was darker due to its electron density. Also, cystocytes generally exhibited larger nuclei and the cytoplasm was always filled with numerous mitochondria. Interestingly, some cystocytes already showed first signs of differentiation since they reached already twice the size of the other cystocytes (Figure 42 A and D). Therefore, the cystocytes within this section of the germarium can be differentiated into pro-oocytes, cells with the potential of differentiating into the oocyte, and pro-trophocytes, which will eventually develop into nurse cells (trophocytes). Nevertheless, pro-oocytes and pro-trophocytes were never observed to be infected with the bacterial endosymbionts this early

during oogenesis, while the somatic cells in the outer layers of the lower germarium as well as somatic cells in the inner layers of each ovariole became progressively infected (Figure 42 A, E and F).
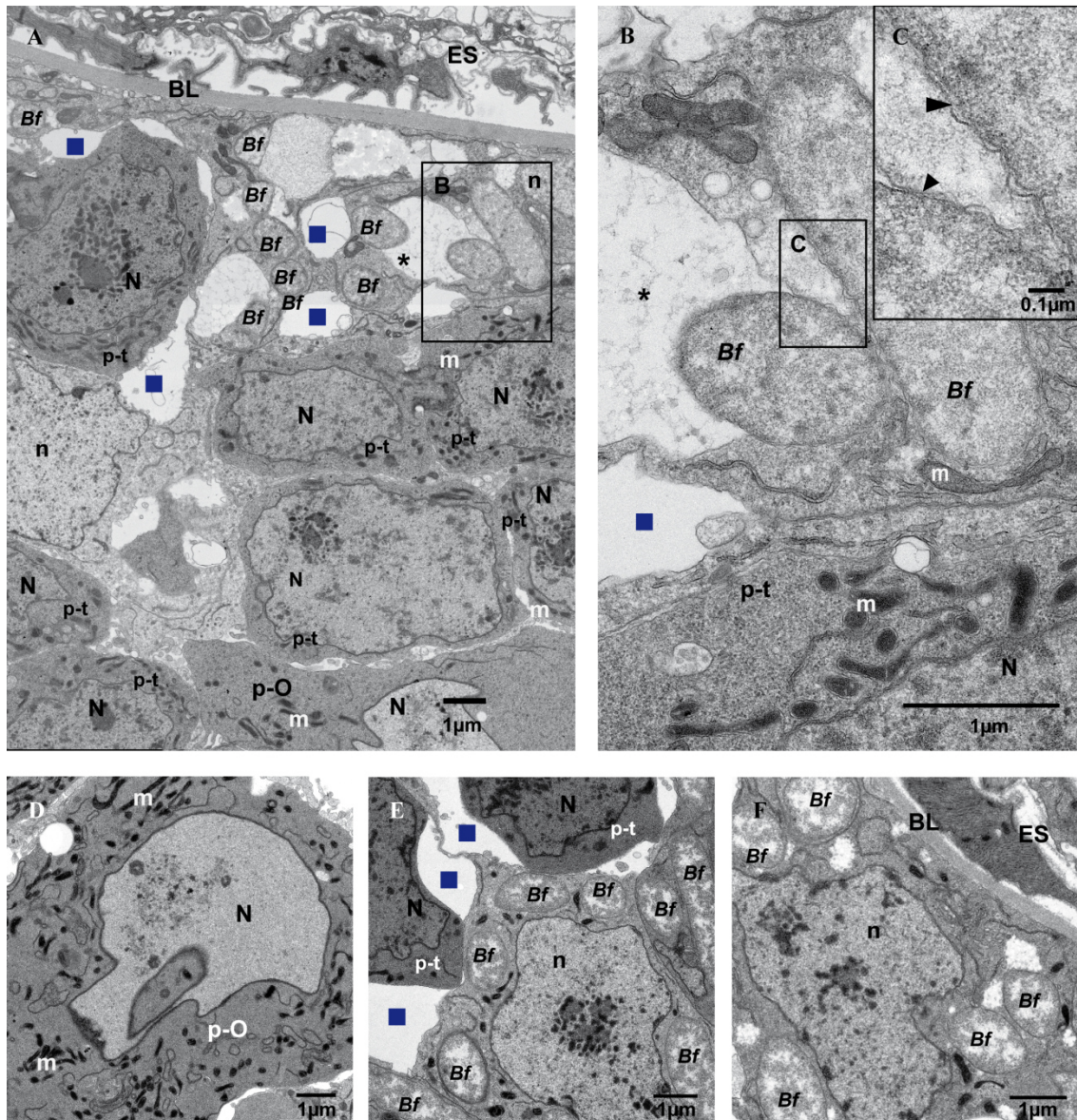

Figure 42 The morphology and ultrastructure of *C. floridanus* ovarioles within the "early infection zone" of the lower germarium presented in electron micrographs of longitudinal tissue sections. **(A)** *B. floridanus* (*Bf*) were observed in small somatic cells lying under the basal lamina (BL) and the ovariolar epithelial sheath (ES), which surround each ovariole. The cystocytes show characteristically large nuclei and a dark cytoplasm. These cells which can already be classified into pro-trophocytes (p-t) and pro-oocytes (p-O) remained uninfected **(A)** and **(D)**, whereas small somatic cells harboured the endosymbionts **(A)**, **(E)** and **(F)**. **(B)** The enlargement of (A) shows a somatic cell containing mitochondria (m) and *Bf* residing free in the cytoplasm or in a vacuole-like structure (marked with asterisk). **(C)** Furthermore, the enlargement of B) presents the characteristic cell wall structure of Gram-negative bacteria (pointed arrowheads). Holes in the tissue sections which are likely to be artefacts of fixation and Epon embedding are marked with blue squares. (N = nuclei of cystocytes; n = nuclei of somatic cells)

Following the anterior-posterior axis of the ovariole, two crucial events take place in the lower germarium. As the bacterial endosymbionts were observed to spread throughout the whole tissue, virtually every cell of somatic origin (CoSO) appeared to become infected with *B. floridanus*. These somatic cells will develop into follicle cells forming a huge mass in which the germ cells are

embedded in. Furthermore, one pro-oocyte of a 32-cell cyst differentiates into the determined oocyte indicated by the faster growth rate when compared to the nurse cells (Figure 43). Additionally, the nuclei of determined oocytes exhibit an irregular morphology as characteristic DNA-free accessory nuclei are generated (Figure 43 C-E).
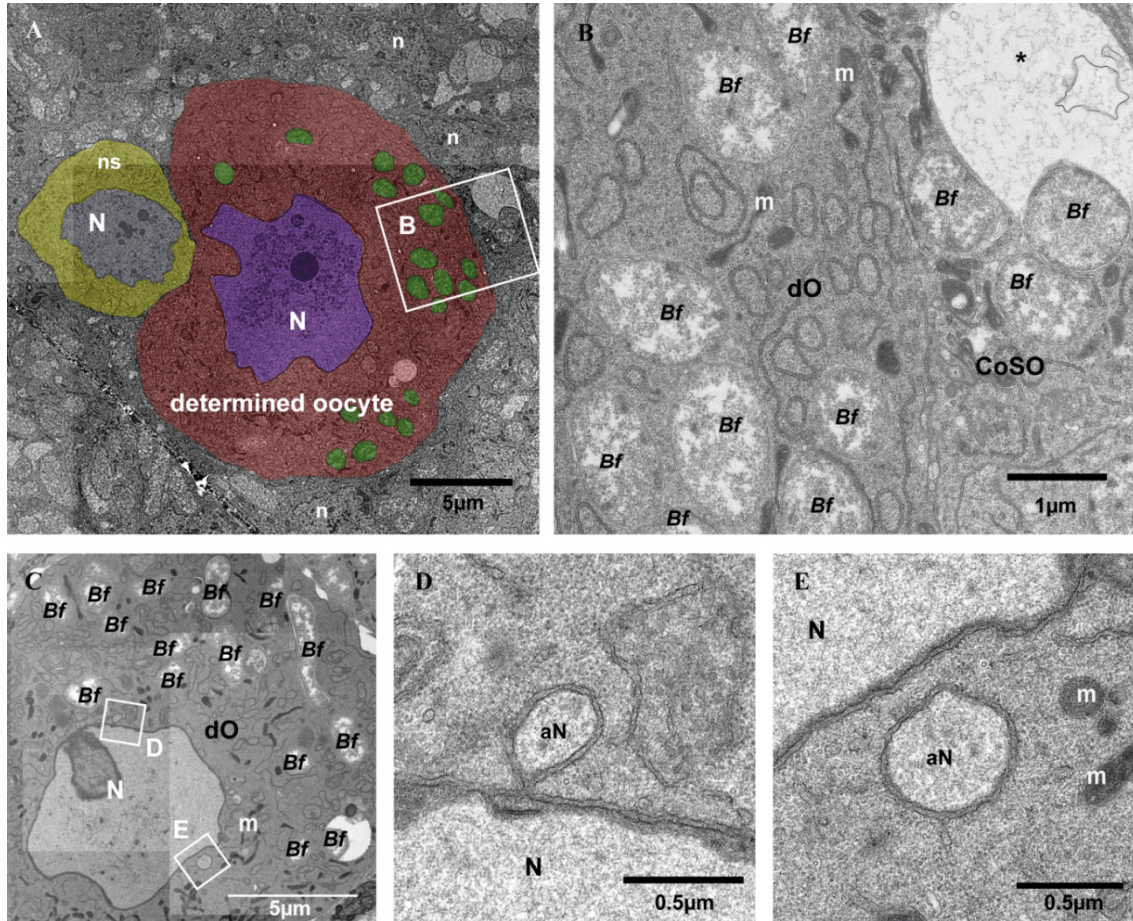
Figure 43 Ultrastructure of *Blochmannia* infected determined oocytes presented in electron micrographs of longitudinal sections. **(A)** Due to progressing differentiation, the determined oocyte (dO) is larger than the nurse cell (nc), and the dO's nucleus shows an irregular morphology. The cytoplasm of the dO (annotated in red) harbours *B. floridanus* (*Bf*; annotated in green). The endosymbionts were observed exclusively in the cytoplasm of the determined oocyte and the surrounding cells of somatic origin (CoSO) but not in the cytoplasm of neighbouring nurse cells (cytoplasm annotated in yellow). **(B)** The enlargement of (A) shows the cytoplasm of the dO filled with mitochondria (m) and *B. floridanus*. *Bf* also reside within vacuole-like structures (marked with asterisk). **(C)-(E)** The formation of DNA-free accessory nuclei (aN) indicates the next step in oogenesis. The determined oocytes appeared to become initially infected with *B. floridanus* in step with the beginning of aN formation. (N = nuclei of nurse cells or determined oocytes; n = nuclei of CoSOs)

At this stage of oogenesis, *B. floridanus* was detected inside the cytoplasm of these very young determined oocytes which is in concert to observations of Blochmann (1882) and Buchner (1918) who described the initial infection of young oocytes in *C. ligniperdus* occurring in step with the formation of accessory nuclei (Blochmann, 1882; Buchner, 1918). As the determined oocytes are surrounded by the CoSOs it can be assumed that the bacteria are transmitted over the whole surface of the oocyte rather than at one particular pole. Astonishingly, the endosymbionts appeared to be exclusively transmitted into the determined oocytes, as *B. floridanus* was never detected in the cytoplasm of nurse cells.
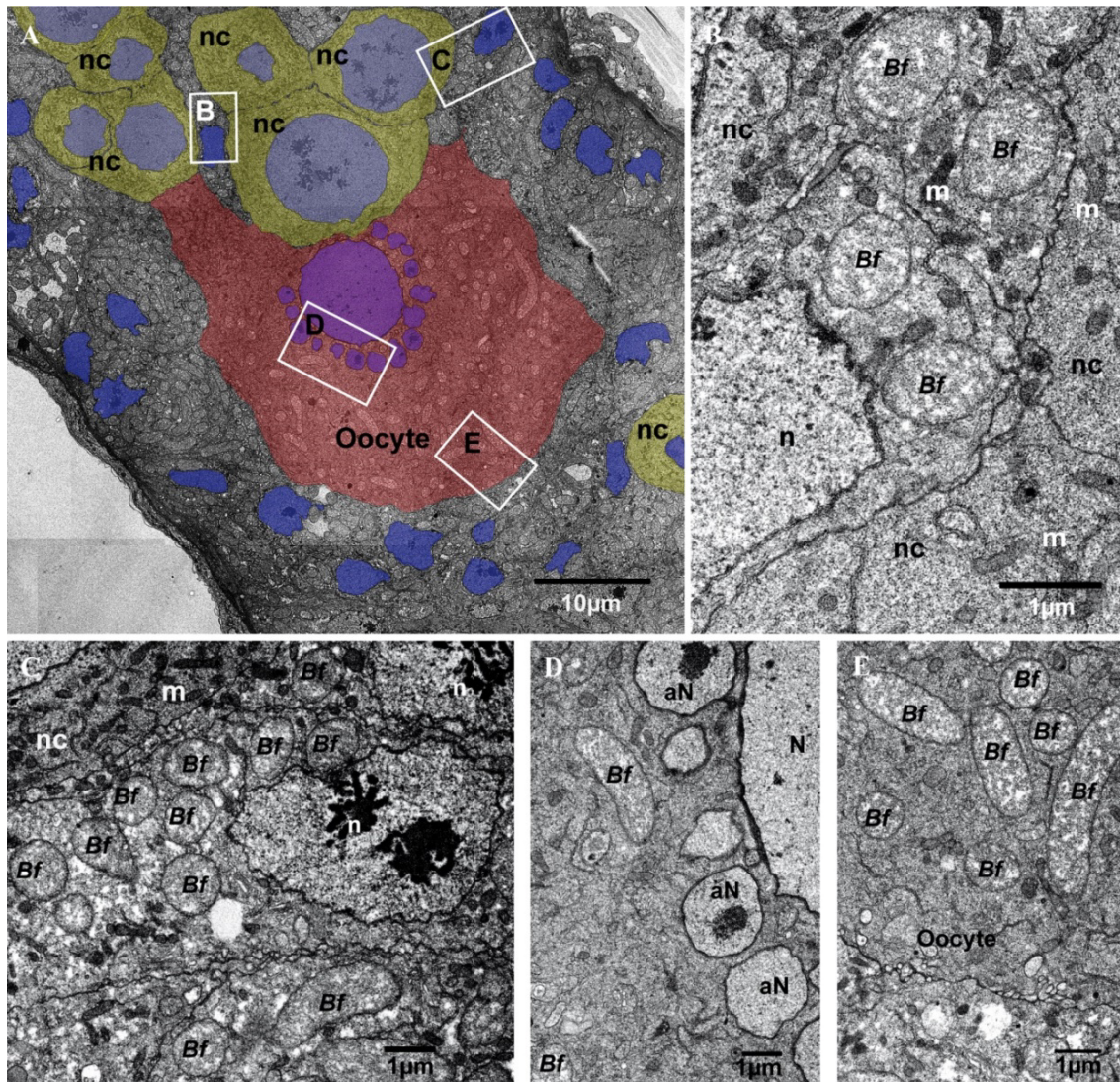
Figure 44 Electron micrographs of longitudinal sections showing the ultrastructure of an ovariole segment presenting the beginning of egg-chamber formation. **(A)** The cytoplasm of the oocyte (annotated in red) harbours bacterial endosymbionts, whereas there are no bacteria in the nurse cell (nc) cytoplasm (annotated in yellow). The somatic cells which differentiate into follicle cells (nuclei (n) annotated in blue) are distributed around the oocyte but also intercalated between the nurse cells. **(B)** and **(C)** The enlargements show somatic/ follicle cells between the nurse cells (B) and around the oocyte (C). *B. floridanus* (*Bf*) reside within their cytoplasm. **(D)** The enlargement of A) shows the nucleus (N) of the oocyte exhibiting typical accessory nuclei (aN). Scattered *Bf* reside in the ooplasm at this stage of egg-chamber formation.

In the last segment of the basal germarium egg-chamber formation completes the chain of events concerning germ cell differentiation (Figure 44 and Figure 45). The three major hallmarks of the formation of egg-chambers (follicles) before entering the vitellarium are (i) the establishment of compact nurse cell packages; (ii) the eventual differentiation of somatic cells into follicle cells, which will form a single cell layer around the oocyte; and (iii) the progressing transmission of *B. floridanus* into the oocyte. Whilst in the beginning of egg-chamber formation, a considerable number of somatic cells were still interspersed among the nurse cells (Figure 44 A and B), these cells differentiating into follicle cells gradually surrounded the oocyte, separating the latter from compact nurse cell packages without such intercalated somatic cells (Figure 45 A and D). As mentioned before, the cytoplasm of the nurse cells remained symbiont-free throughout the whole oogenesis. Instead, the endosymbionts

were continuously transmitted from the follicle cells into the growing oocytes without even passaging through the nurse cells.
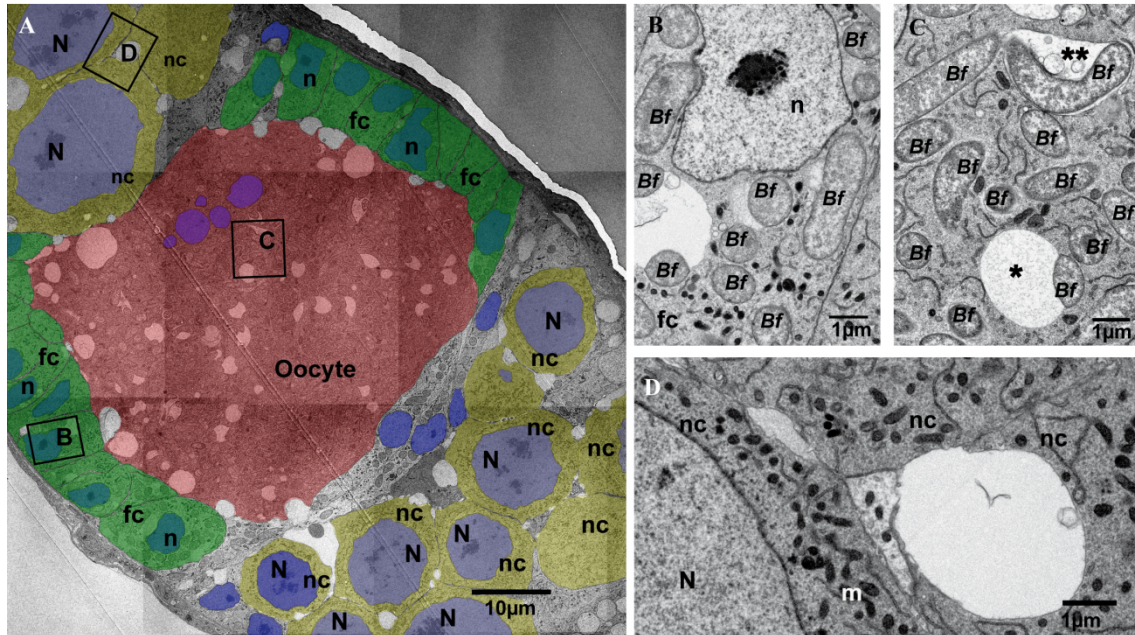

Figure 45 Electron micrographs presenting the ultrastructure of an almost completely formed egg chamber in a *C. floridanus* ovariole. **(A)** The cytoplasm of the oocyte (annotated in red) harbours *B. floridanus* (*Bf*), and most of the follicle cells (fc; annotated in green) are distributed in a single cell layer around the oocyte. The nurse cell cytoplasm is annotated in yellow. **(B)** The enlargement of a follicle cells shows that their cytoplasm harbours numerous endosymbionts at this stage of oogenesis. **(C)** The endosymbionts (*Bf*) reside in the ooplasm or inside of vacuole-like structures with a single membrane (marked with single asterisk) or a double membrane (marked with two asterisks). **(D)** The enlargement of the nurse cell (nc) package shows the *Blochmannia*-free cytoplasm of nurse cells. (m = mitochondria; N = nuclei of nurse cells; n= nuclei of follicle cells)

During progressing egg-chamber formation and oocyte growth within the vitellarium, the amount of *Blochmannia* in follicle cells and the oocytes changed in an opposed manner as already observed via FISH (chapter 4.5.1). While the number of *B. floridanus* in follicle cells slowly decreased, the ooplasm appeared to be increasingly packed with the endosymbionts. In inchoate egg-chambers, only scattered bacteria were observed in the ooplasm leaving the larger proportion of the latter bacteria-free (Figure 44). The ooplasm of growing oocytes became increasingly occupied by *Blochmannia*, and only the ooplasm around the accessory nuclei remained less packed with the endosymbionts (Figure 45 and Figure 46). It is worth mentioning that *B. floridanus* was either observed free in the ooplasm or within vacuole-like structures with a single membrane or a double membrane (Figure 45 B). Finally, follicle cells which then appeared bacteria-free surround the oocyte completely (Figure 46), and the complete egg-chambers undergo oocyte growth and later on vitellogenesis within the vitellarium. In this stage of oogenesis, the division and massive elongation of *B. floridanus* was observed, and the bacteria reached lengths of up to 8 µm (Figure 46 B and C).
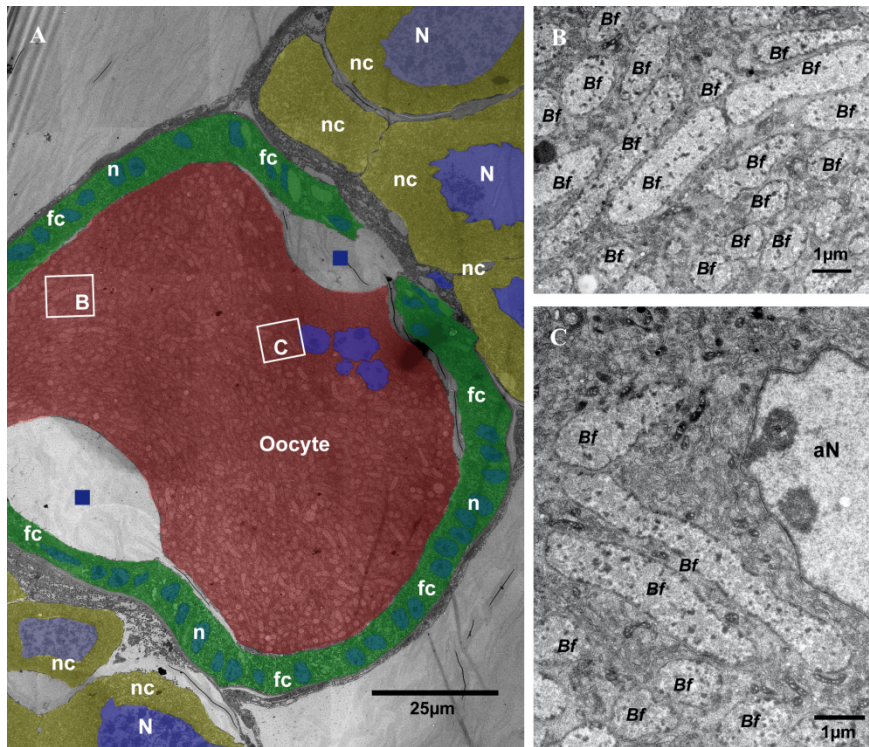
Figure 46 The *C. floridanus* egg-chamber shortly before vitellogenesis. The electron micrographs of longitudinal sections of an ovariole show **(A)** the cytoplasm of the oocyte (annotated in red) densely packed with *B. floridanus* (*Bf*). The follicle cells (fc; green) build a single cell epithelium around the oocyte which is nestled in nurse cell packages (nc, annotated in yellow). **(B)** and **(C)** The enlargements of (A) show *Bf*, which can only be found in the ooplasm with the exception of the region close to the accessory nuclei (aN). *Bf* elongates to up to 8 µm and mostly resides free in the ooplasm. Holes in the tissue sections which are likely to be artefacts of fixation and Epon embedding are marked with blue squares. (N = nuclei of nurse cells; n = nuclei of follicle cells)

However, the ratio of *Blochmannia*-filled and *Blochmannia*-free ooplasm changed drastically again during further oocyte growth. Due to the storage of enormous amounts of protein and lipid yolk platelets during vitellogenesis, the 50-100 µm long oocytes packed with *B. floridanus* (Figure 46) reached a length of up to 0.5-1.0 mm and the bacteria appeared to be shifted towards the posterior pole of the fully grown oocyte which will eventually enter the oviduct (Figure 47).
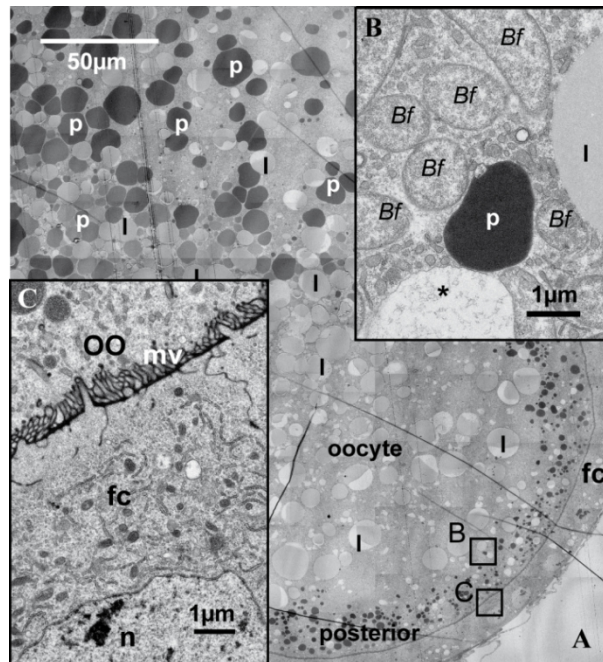
Figure 47 The ultrastructure of the posterior pole of a fully grown oocyte presented in a longitudinal section. **(A)** The overview shows the cytoplasm of the fully grown oocyte which is filled with yolk platelets (protein yolk (p) and lipid yolk (l)). A single layer of follicle cells (fc) surrounds the oocyte (OO). **(B)** *B. floridanus* (*Bf*) was found at the posterior pole of the OO residing either free in the ooplasm or rarely in vacuole-like structures (marked with asterisk). **(C)** The enlargement of (A) shows the bacteria free cytoplasm of follicle cells. Numerous microvilli (mv) line the boundary between follicle cells and the oocyte. (n = nucleus of a follicle cell)

4.5.3 The intracellular localisation of *Blochmannia floridanus*

Previously published data indicate that *B. floridanus* most frequently resides free within the cytoplasm of specialised bacteriocytes within the midgut tissue and in fully grown oocytes (Sauer et al., 2002; Schröder et al., 1996). The extensive analysis of the ovarian tissue of *C. floridanus* workers confirmed these data, as most of the endosymbionts were observed residing in the cytoplasm of the oocytes and follicle cells. However, *Blochmannia* was also found inside of vacuole-like structures (marked with asterisks in Figure 42 - Figure 47). The abundance of these vacuoles increased during the development of determined oocytes into growing oocytes within egg-chambers. Whereas the number of such vacuole-like structures was quite low in the determined oocytes and very young oocytes, the number of vacuoles continuously increased and reached a maximum in young egg-chambers when the endosymbionts accumulated massively in the oocytes. The amount of vacuole-like structures observed decreased again when follicle cells were bacteria free, and *B. floridanus* appeared to be shifted towards the posterior pole of the oocyte during vitellogenesis. Therefore, bacteria-containing vacuoles were observed only rarely in the ooplasm of fully grown oocytes in the late stages of oogenesis (Figure 47).

The detailed analysis presented in this work enables the discrimination between two types of vacuoles. Most of the observed vacuole-like structures exhibit a single membrane which encloses one or more bacteria (Figure 48 A and G; marked with single asterisks; membrane marked with pointed arrowheads). Additionally, a tomogram and the respective 3-D model of *B. floridanus* residing within

such a vacuole-like structure were generated. The tilt series for tomogram generation were recorded from at least -65° to +65° with 1° increments using the software SerialEM. Nine tilt series were further processed following the processing steps of the toolset eTOMO within the software package IMOD v4.7. Finally, tomograms were generated using the option "Back Projection", and tomograms were joined in z-axis orientation in order to create the full tomogram. Subsequently, membranes of bacteria and the vacuole, as well as mitochondria were annotated using the 3dmod tool of the software IMOD, and the 3-D model was generated. Interestingly, even though the vacuole membranes appeared to tightly envelop *B. floridanus*, the vacuoles reached considerable length when more than one bacterium was observed to reside within the vacuole (Figure 48 J and K). The second type of these vacuole-like structures exhibited a double membrane surrounding *B. floridanus* (Figure 48 C, E and G; marked with double asterisks; membranes marked with stealth arrowheads). Additionally, these double membrane vacuoles frequently comprised small vesicles which were only rarely observed in single membrane vacuoles (Figure 48 C and G; marked with red squares).

In parallel with the massive abundance of *B. floridanus* inside of these vacuoles in oocytes of young egg-chambers, another interesting phenotype was observed at the delimitative cytoplasm membrane between follicle cells and the oocytes (Figure 48 L-N, marked with red asterisks). Numerous widenings of the intercellular space between follicle cells and the oocyte formed invagination-like structures towards the ooplasm. These invaginations were at least 0.5-1.0 µm wide and were generally several micrometres long. Furthermore, all invaginations observed contained small vesicles which might indicate transport processes between follicle cells and the oocyte, as it was demonstrated for the ovarian tissue of *Apis mellifera* (Fleig, 1995).

The formation of these vacuoles around *B. floridanus* and the observation of such invagination-like structures may be "snap-shots" of the transport mechanisms of the bacteria from follicle cells into the oocytes. However, as mentioned before, the dissection of ovarioles itself might have caused mechanical stresses and therefore might have led to disruptions and holes within tissue sections of the germarium (Figure 41). Additionally, several holes were observed in sections of the vitellarium (Figure 45 and Figure 46), and it can be argued that the observed invagination-like structures were artefacts rather than naturally occurring phenomena (Figure 48). In order to reduce artefacts in the samples and to generally improve the quality of tissue fixation, ovarioles were dissected and applied for high pressure freezing which allows the preservation of tissues up to a thickness of 0.6 mm.
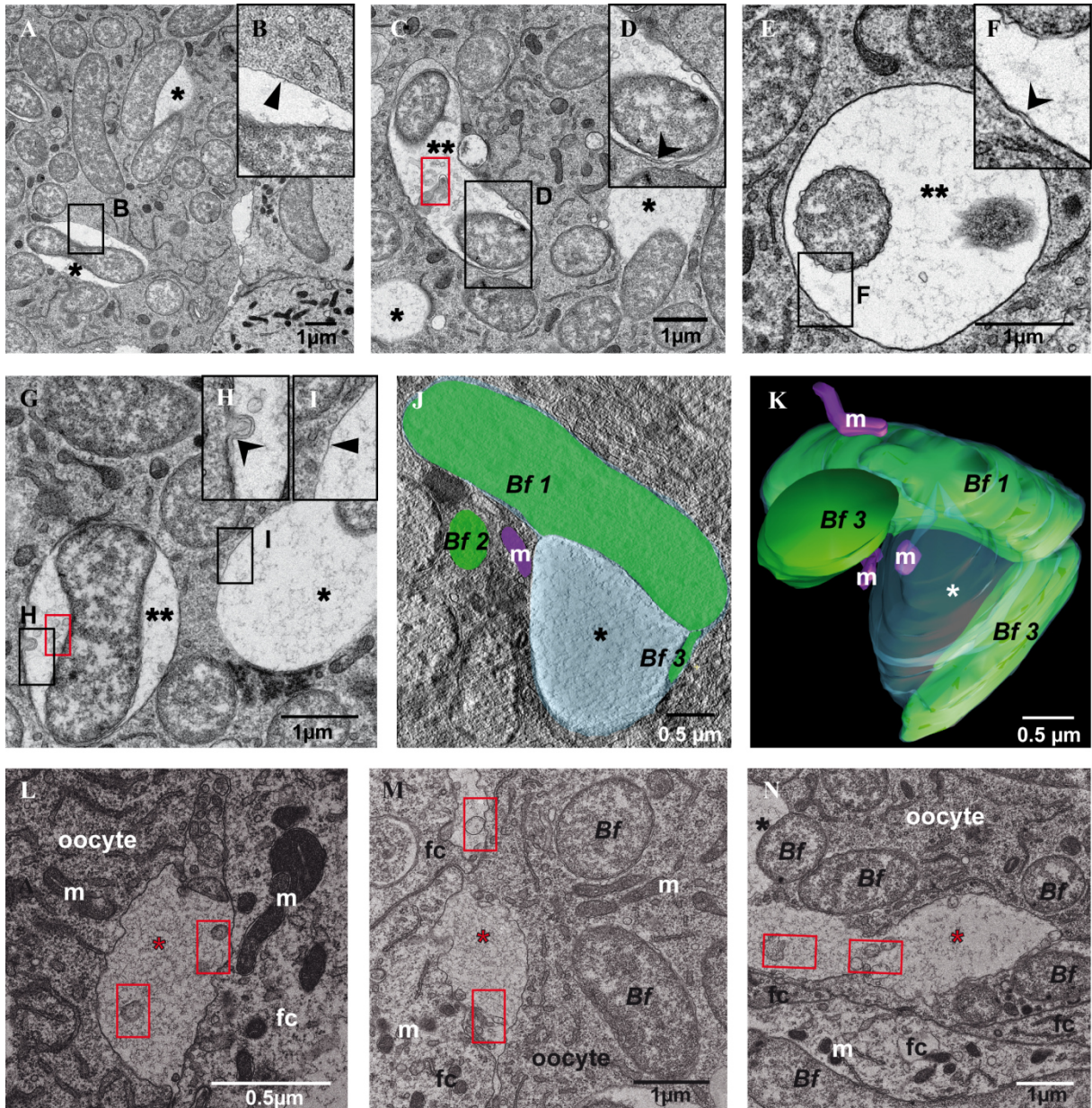
Figure 48 The intracellular localisation of *B. floridanus* within the oocytes. **(A) - (K)** The electron micrographs show *B. floridanus* (*Bf*) residing either freely in the cytoplasm or within vacuole-like structures inside oocytes within young egg-chambers. **(A)**, **(G)** and **(J)** Most of the vacuoles exhibit a single membrane (marked with single asterisk) which encloses one or more bacterial endosymbionts. **(B)** and **(I)** Single membranes are marked with pointed arrowheads in the enlargements of (A) and (G). **(J)** The picture of a tomogram slice shows *B. floridanus* (annotated in green) inside a vacuole (asterisk). The vacuole (light blue) exhibits a single membrane which completely surrounds the bacterial endosymbiont. **(K)** The annotation of several structures within the tomogram resulted in the generation of the 3-D model showing two endosymbionts inside the single membrane vacuole. The bacterial cell wall is annotated in green, the vacuole membrane is annotated in light blue, and mitochondria within the cytoplasm are annotated in purple. **(C) - (H)** A second type of vacuoles comprises double membranes (double black asterisks) Membranes are marked with stealth arrowhead in the enlargements of the structures in (D), (F) and (H). Additionally, small vesicles can be observed in double membrane vacuoles (marked with red squares). **(L) - (N)** Furthermore, widenings of the intercellular space (red asterisks) can be observed at the cytoplasmic membrane of the follicle cell and oocyte. These "gaps" form invagination-like structures towards the ooplasm and are at least 0.5-1.0 µm wide or even longer. All invaginations observed contain small vesicles (marked with red squares). (m = mitochondria; fc = follicle cell)

143

In a trial experiment, especially thinner parts of the ovarioles, comprising the apical and basal germarium including *B. floridanus*, were fixed via high pressure freezing (HPF). However, the whole tissue analysed showed signs of severe ice crystal formation which generally occurs when the tissue is not fixed fast enough (Figure 49). Thus, numerous small ruptures destroyed the integrity of cellular structures such as the cytoplasm itself, mitochondria and membranes. Even structures in the outer layers of the ovariole were fixed improperly as indicated by disrupted cytoplasmic membranes and nuclear envelops in the pro-trophocytes and especially the Gram-negative cell wall structure of *B. floridanus*. Furthermore, the extent of the tissue damage increased towards the inner layers of the ovariole. It can be assumed that HPF as applied in this trial experiment would not be sufficient for fixation of even thicker parts of the ovarioles such as the vitellarium. Therefore, it was not possible to distinguish suitable conditions to improve the fixation via HPF and to further investigate the nature of vacuole-like structures and invaginations within young egg-chambers in the vitellarium.
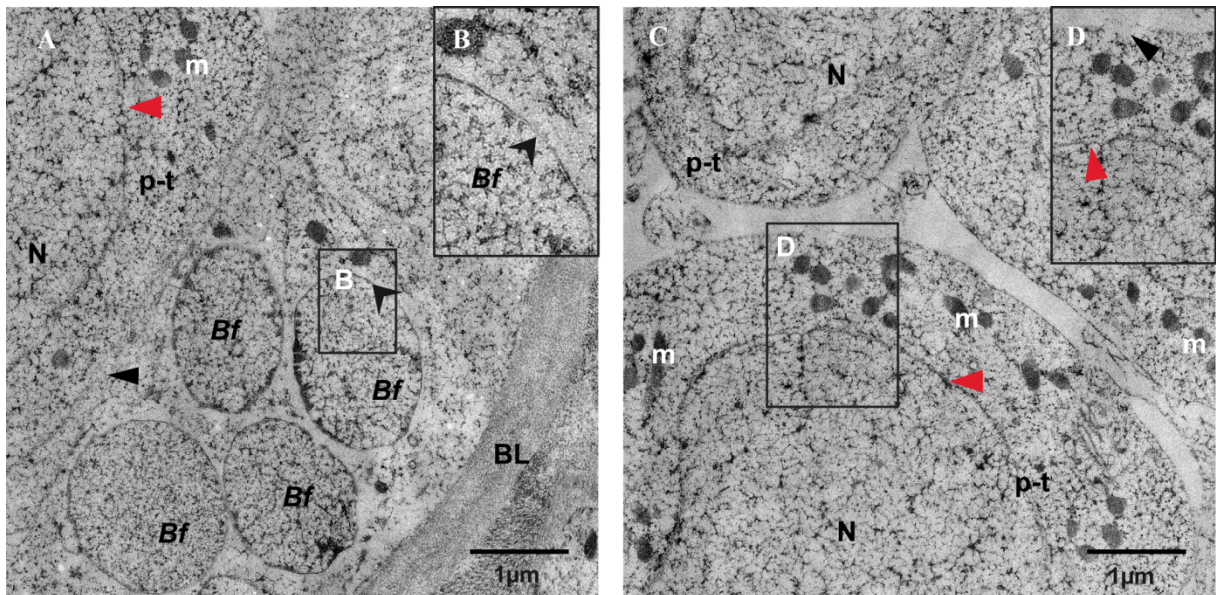
Figure 49 Ovarian tissue of *C. floridanus* fixed via high pressure freezing. Ovarioles were dissected, and several sections comprising the apical and lower germarium were applied for HPF. **(A)** Though the germarium is comparatively thin, the whole tissue shows signs of severe ice crystal formation. Even structures in the outer layers of the ovariole are fixed improperly as indicated by disrupted cytoplasmic membranes (black arrowhead) and nuclear envelops (red arrowhead) in the pro-trophocytes (p-t). **(B)** Also, the bacterial cell wall structures are not completely preserved (marked with stealth arrowhead), even though the endosymbionts (*Bf*) reside directly under the basal lamina (BL) at this stage of oogenesis. **(C)** The extend of the tissue damage increased towards the inner layers of the ovariole. Numerous small ruptures destroy the cellular integrity of the pro-trophocytes as well as the integrity of the whole tissue. **(D)** Disrupted cytoplasmic membranes (black arrowhead) and nuclear envelops (red arrowhead) are shown in an enlargement of (C).

4.5.4   The distribution of *Blochmannia floridanus* in midguts and ovaries of rifampicin treated ants

Previous work has shown that antibiotic treatment of *C. floridanus* workers with either rifampicin or tetracycline resulted in the nearly complete clearance of *B. floridanus* from the bacteriocytes within the midgut (Sauer et al., 2002; Zientz et al., 2006). Interestingly, when provided with fresh larvae of their mother colony, subcolonies treated with antibiotics raised significantly fewer pupae from the larvae than untreated subcolonies. Though neither possible toxic effects of the antibiotics nor the loss of the endosymbionts within the midgut tissue led to significantly higher levels of mortality of the workers, their fitness in terms of successfully raising pupae seemed to be impaired (Zientz et al., 2006).

In this work, a trial experiment was performed in order to observe the influence of antibiotic treatment on the ovariolar distribution of *B. floridanus*. An orphaned *C. floridanus* colony (C90) with egg-laying workers and successfully raised males was equally divided. Males, eggs, larvae and pupae were removed from the subcolonies. While one group (45 major workers and 60 minors) was treated with rifampicin, the other group was fed normally over a period of eight weeks. The number of adult workers in the group treated with antibiotics decreased massively during the period of observation. While only 16 adults (minors and majors) were counted within the treated subcolony after eight weeks, 39 adults survived in the control group. Moreover, about 50 eggs and larvae, 30 pupae (P1-P3) and one male were raised only within the control group, whereas only 10 very small larvae (L1) but neither pupae nor adult males were found in the rifampicin treated group. In order to investigate the distribution of *B. floridanus* in the ovaries and midguts after antibiotic treatment, tissues were dissected and applied for the FISH procedure as described above (chapter 3.11.1).

Within the ovarioles of the control animals *B. floridanus* appeared to be distributed as described before. The apical germarium was bacteria-free, and the basal germarium showed massive infection with the endosymbionts. The vitellarium of such an ovariole exhibited the typical alternation of growing oocytes filled with *B. floridanus* and bacteria-free nurse cell packages (Figure 50 A and B). Furthermore, the bacteriocytes within the midgut tissue of workers with active ovaries were massively filled with the endosymbionts (Figure 50 A and B) which confirms previous observations (Stoll et al., 2010).

Interestingly, the treatment with rifampicin over a period of 8 weeks did not result in complete clearance of *B. floridanus* from the midgut tissue of the analysed workers. In particular, almost equal numbers of bacteriocytes filled with endosymbionts were detected when compared to the amount of bacteria-filled bacteriocytes in midguts from control animals (Figure 50 C and D). However, the ovarian tissues of workers fed with rifampicin were completely cleared from the endosymbionts. *B. floridanus* was neither observed in the basal germarium of ovarioles nor in the oocytes within the vitellarium (Figure 50 C and D). Therefore, the period of eight weeks for rifampicin treatment was

sufficient for the removal of endosymbiotic bacteria from active ovaries in *C. floridanus* workers. Furthermore, signs of tissue damage were not observed as the overall structure of ovaries (length and number of ovarioles per ovary; number and appearance of oocytes) was similar to ovaries in control animals.
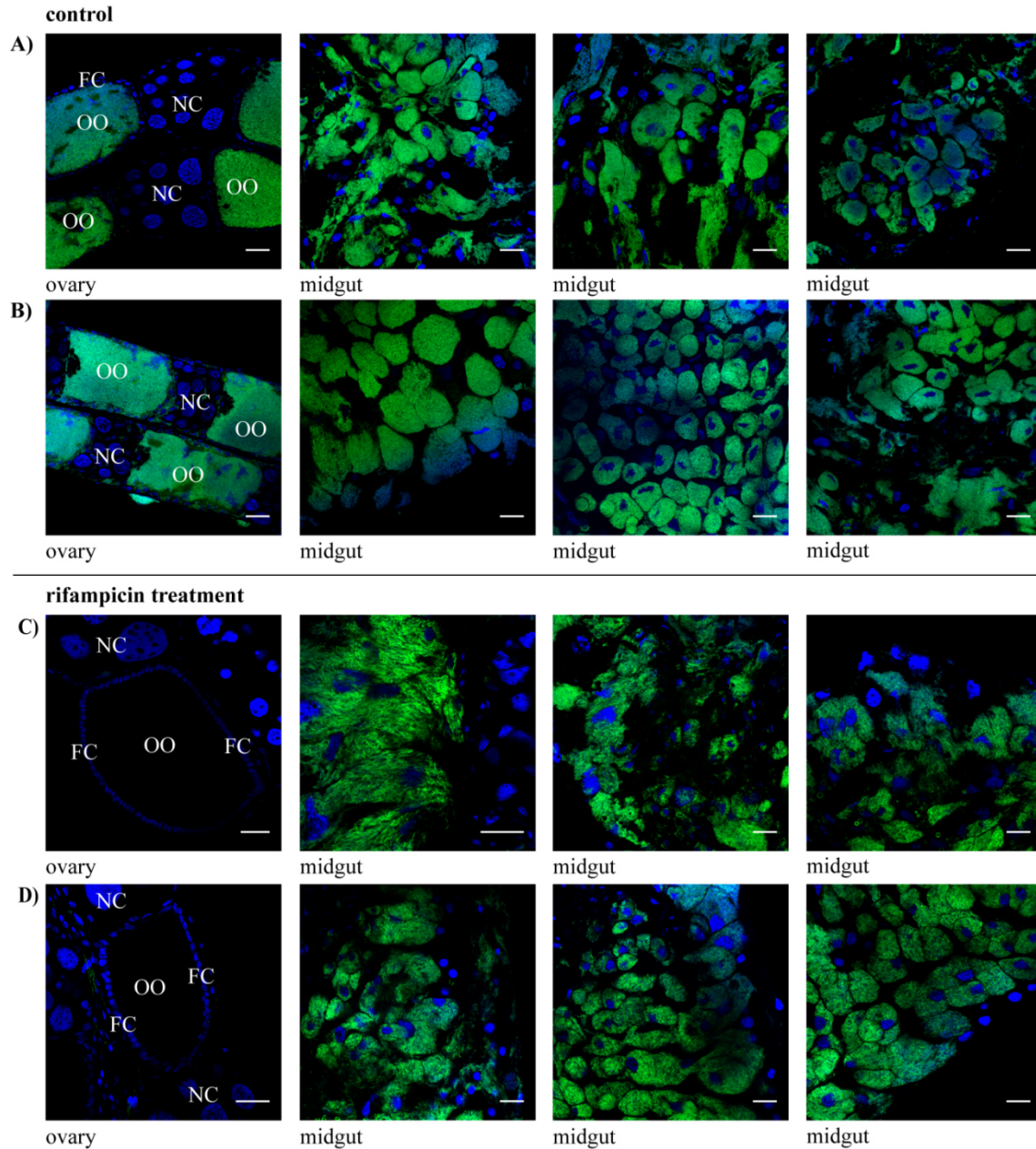
Figure 50 Antibiotic treatment of *C. floridanus* and its effects on the distribution of *B. floridanus* in ovaries and midguts of the ants. Queenless workers of *C. floridanus* were either treated with rifampicin or provided with the normal diet. The distribution of the endosymbionts in the ovaries and midguts was analysed. For optimal illustration, fluorescence images were taken of the vitellarium of each analysed ovariole. For workers with active ovaries, three sections representing different positions within the midgut are shown. **(A)** and **(B)** *B. floridanus* (Bfl172_Alexa488, green) was detected within the oocytes (OO) in the vitellarium of the dissected ovarioles of two control workers. Additionally, bacteriocytes in the midgut tissue of these workers appeared to be filled with the endosymbionts. **(C)** and **(D)** Rifampicin treated ovarian tissue of workers appeared to be bacteria-free, and *B. floridanus* was not detected in growing oocytes. However, the midgut tissue of these workers still contained high numbers of bacteriocytes filled with the endosymbionts. Scale bars represent 25 µm. DNA was counterstained with DAPI (blue).

In addition to the qualitative analysis of ovarian tissue, midguts of male larvae either originated and raised by workers, which were treated with rifampicin, or raised by untreated ants were dissected. The bacteriocytes in the midgut tissue of males originated and raised by the latter appeared to be filled with the endosymbionts (Figure 51 A and B). Interestingly, only very few endosymbionts were observed in the midguts of male larvae which developed from eggs laid by workers previously fed with the antibiotic. The majority of the midgut tissue appeared to be bacteria-free (Figure 51 C and D).
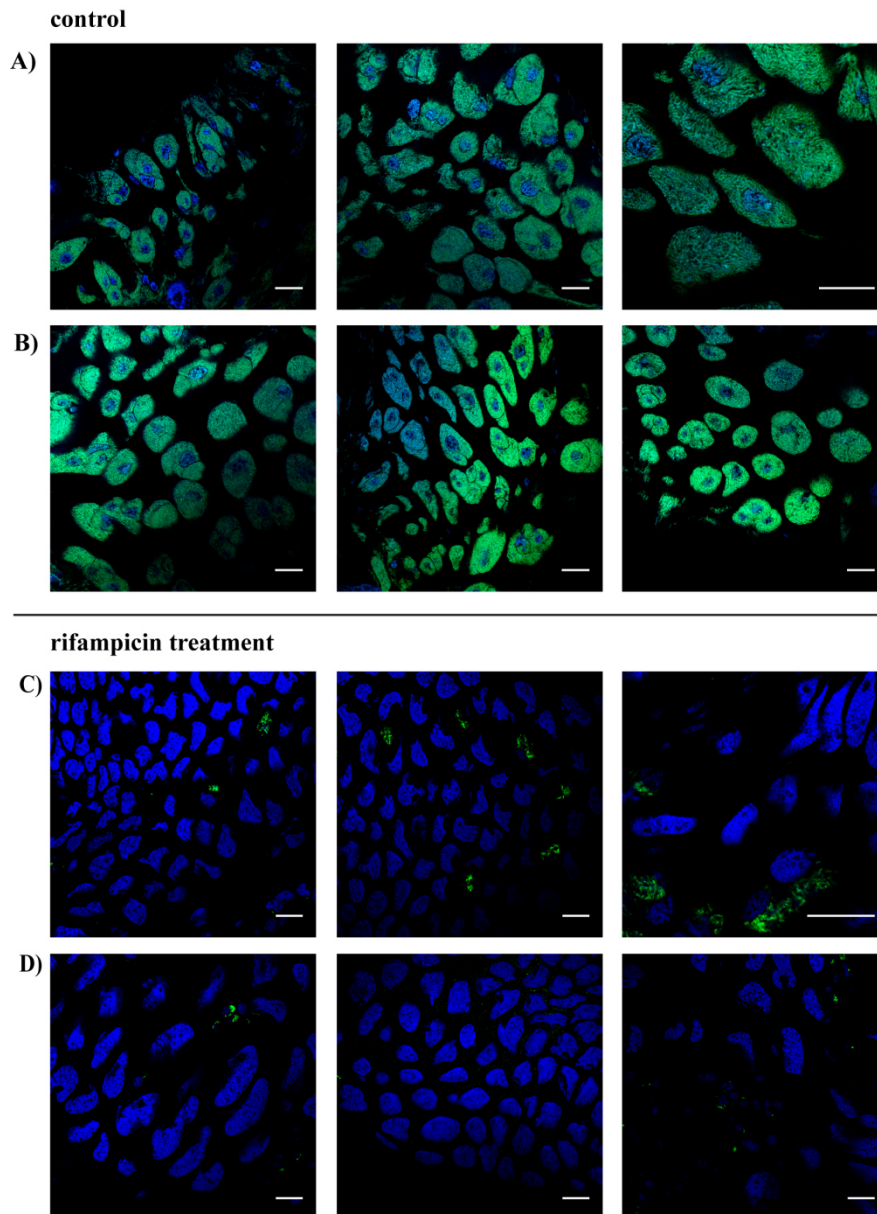
Figure 51 Antibiotic treatment of *C. floridanus* workers and its effects on the distribution of *B. floridanus* in midgut tissue of *C. floridanus* male larvae. Queenless workers of *C. floridanus* were either treated with rifampicin or provided with the normal diet. The distribution of the endosymbionts in the midguts of male larvae was analysed. Fluorescence images of three different sections of the tissue are shown for each dissected midgut. **(A)** and **(B)** High numbers of *B. floridanus* (Bfl172_Alexa488, green) were detected within bacteriocytes of male larvae raised by untreated female workers. **(C)** and **(D)** Significantly less endosymbionts were detected in the midgut tissue of male larvae raised by workers which were treated with the antibiotic. In some sections the midgut tissue appeared to be almost bacteria-free. Scale bars represent 25 μm and DNA was counterstained with DAPI (blue).

Taken together, the low number of male larvae in the rifampicin treated subcolony and the extremely low amount of endosymbionts observed within the larval midguts suggest that the rifampicin-mediated removal of *B. floridanus* from the ovarian tissue of workers results in oocytes and therefore eggs without or with significantly less endosymbionts. Hence, only few male larvae and only a small amount of *B. floridanus* within their midguts were observed. Within the tested subcolony male larvae were never observed to be raised further and did not complete ontogeny.

4.6 The expression of putative immune genes in the ovarian tissue of *C. floridanus*

In order to analyse the expression of putative immune genes in ovarian tissue of *C. floridanus*, ovaries, midguts and residual body parts were dissected for subsequent RNA purification. To avoid the detection of constitutively expressed immune genes, the fat body, which is the major immune organ in the ants, was not included in the residual body samples. Differential expression of genes with putative immune response was analysed via qRT-PCR and gene expression was compared between the two symbiont bearing organs, the midgut and ovary, and the residual body, which does not harbour *B. floridanus*.

Before the genes of interest were tested for their expression in different tissues of the ants, an appropriate reference gene of *C. floridanus* for normalisation of the qRT-PCR data was identified. Four supposedly housekeeping genes, *ribosomal protein L32* (*rpL32*, GenBank Acc. No.: EFN68969), *rpL18* (GenBank Acc. No.: EFN68908), *elongation factor 1-alpha* (*EF1α*, GenBank Acc. No.: EFN72500) and *glyceraldehyde-3-phosphate dehydrogenase* (*GAPDH*, GenBank Acc. No.: EFN69158), were analysed with respect to stability of expression levels (Ct-values) of the genes in all tissue samples. Hence, Ct-values of the genes were compared for all cDNA samples derived from different tissues (ovary, midgut and residual body) obtained from animals of two different colonies (C90 and C152) using the BestKeeper (BK) software tool (Pfaffl et al., 2004). Within this tool, the standard deviation-value (SD-value) is the most important criterion for evaluating the stability of reference genes. The expressions of all reference genes tested in this study showed low Ct variations (0.40 < SD [± Ct] < 0.87) as well as low regulation (1.32 < SD [± x-fold] < 1.82). Therefore, all genes were well suited as reference genes (Table 18). Nevertheless, *rpl32* (SD [± Ct] = 0.40; SD [± x-fold] = 1.32) was chosen as an internal standard for target gene expression analysis, as the values determined with the BK software tool were most similar to the BestKeeper index (SD [± Ct] = 0.55; SD [± x-fold] = 1.47).

Table 18 Statistical analyses of four candidate reference genes, *EF1a*, *rpL32*, *rpL18* and *GAPDH*, based on their Ct-values. Results were obtained utilising the BestKeeper software tool. In the last column, all standard candidates are combined in the BestKeeper (BK) index.

	EF1a	*rpL32*	*rpL18*	*GAPDH*	**BestKeeper**
n	62	62	62	62	62
geo Mean [Ct]	18.99	18.87	19.14	19.16	19.04
ar Mean [Ct]	19.02	18.87	19.17	19.18	19.05
min [Ct]	17.20	18.02	16.80	17.44	17.62
max [Ct]	21.59	20.05	21.33	21.36	20.85
std dev [± Ct]	**0.74**	**0.40**	**0.87**	**0.84**	**0.55**
CV [% Ct]	3.87	2.10	4.52	4.39	2.90
min [x-fold]	-3.46	-1.80	-5.06	-3.28	2.68
max [x-fold]	6.04	2.28	4.56	4.60	3.52
std dev [± x-fold]	**1.67**	**1.32**	**1.82**	**1.79**	**1.47**

To obtain a first overview of the expression of putative immune genes in the ovarian tissue, several genes were chosen regarding their putative contribution to recognition and signalling aspects of the immune response: *peptidoglycan recognition protein PGRP-SA* (Cflo_N_g8526t1; previously annotated as PGRP-2; GenBank Acc. No. EFN70060), *PGRP-LC* (Cflo_N_g10272t1; previously annotated as PGRP-LE; EFN63542), *PGRP-LB* (Cflo_N_g103t1; EFN73971), *PGRP-SC2* (Cflo_N_g102t2; EFN73970), *beta-1,3-glucan-binding protein GNBP* (Cflo_N_g5742t1; EFN66519), *GNBP1_3* (Cflo_N_g15215t1; EFN62569), *alpha-2-macroglobulin-like protein TEP2* (Cflo_N_g7345t1; EFN69033), *serine protease persephone* (*psh*; Cflo_N_g8442t1; EFN71579.1), *Toll-interacting protein tollip* (Cflo_N_g363t1; EFN63773) and the *Nuclear factor NF-kappa-B p110 subunit relish* (*rel*; Cflo_N_g6082t1; EFN61437). Statistical analyses for significance of the expression values (dCT) were performed in three steps. At first, the Shapiro-Wilk W Test was performed to generally test for normality. A p-value > 0.05 indicates that the gene expression within a sample is normally distributed (Appendix X, Table 26). With one exception, Ct-values obtained for the tested genes showed a normal distribution in all tissue samples. Only the Ct-values for *relish* were not normally distributed, indicated by a p-value of 0.023. Nonetheless, the general quality of the gene expression levels with respect to distribution of Ct-values was considered suitable for the following statistical analyses. After this initial test, expression levels were tested with an ANOVA in order to investigate the significance of two effects namely "Tissue" and "Colony", as gene expression levels were analysed in three different tissues (residual body, midgut and ovary) and samples were obtained from two different colonies (C90 and C152). "Tissue" was included as fixed effect in the analyses, whereas "Colony" and "tissue nested within colony" ("Tissue(Colony)") were included as random effects (Appendix X, Table 27). Finally, an LSD post hoc test was performed to investigate significant differences of dCt-values between tissue samples (Appendix X, Table 28).

The ANOVA detected the significance of the effect "Tissue" for all genes, which were analysed in this study (*PGRP-SA*: $p < 0.0001$; *PGRP-LC*: $p = 0.0037$; *PGRP-LB*: $p < 0.0001$; *PGRP-SC2*: $p < 0.0001$; *GNBP*: $p < 0.0001$; *GNBP1_3*: $p < 0.0001$; *TEP2*: $p < 0.0001$; *psh*: $p < 0.0001$; *tollip*: $p = 0.0001$; *relish*: $p < 0.0001$). Additionally the interaction "Tissue*Colony" was significant for the following genes: *PGRP-LB* ($p = 0.0011$), *PGRP-LC* ($p = 0.0009$), *GNBP* ($p = 0.0484$), *relish* ($p = 0.0116$) and *TEP2* ($p = 0.0046$). The results revealed that the relative gene expression levels differed significantly between the three analysed tissues. Furthermore, the significance of the interaction "Tissue*Colony" indicated that the expression levels of some genes varied depending on the ant colonies from which the tissue samples originated.

The expression of several genes which were predicted to be involved in the recognition of bacterial or fungal molecular patterns was analysed. Two genes with putative function in the activation of the Toll signalling pathway upon recognition of the lysine-type PGN of Gram-positive bacteria, *PGRP-SA* (Cflo_N_g8526t1) and the *GNBP* (Cflo_N_g5742t1), showed very different expression levels in the

three tissues (Figure 52). The expression of *PGRP-SA* was significantly lower in the midgut (mg) and ovaries (ov) when compared to the relative gene expression in the residual body (rb) (C90: p(rb/mg) < 0.0001, p(rb/ov) < 0.0001; C152: p(rb/mg) < 0.0001, p(rb/ov) < 0.0001). Interestingly, *GNBP* was expressed significantly lower in the midgut when compared to the residual body (C90: p(rb/mg) < 0.0001; C152: p(rb/mg) < 0.0001), but expression levels were higher in the ovaries of *C. floridanus* workers (C90: p(rb/ov) = 0.0588, not significant (ns); C152: p(rb/ov) = 0.009). Another gene encoding a beta-1,3-glucan-binding protein GNBP1_3 (Cflo_N_g15215t1) might be involved in the recognition of either Gram-positive bacteria or fungi. In the midgut tissue *GNBP1_3* was lower expressed, but the gene was more strongly expressed in the ovaries when compared to expression levels in the residual body (C90: p(rb/mg) < 0.0001, p(rb/ov) = 0.0002; C152: p(rb/mg) < 0.0001, p(rb/ov) = 0.0023) (Figure 52).
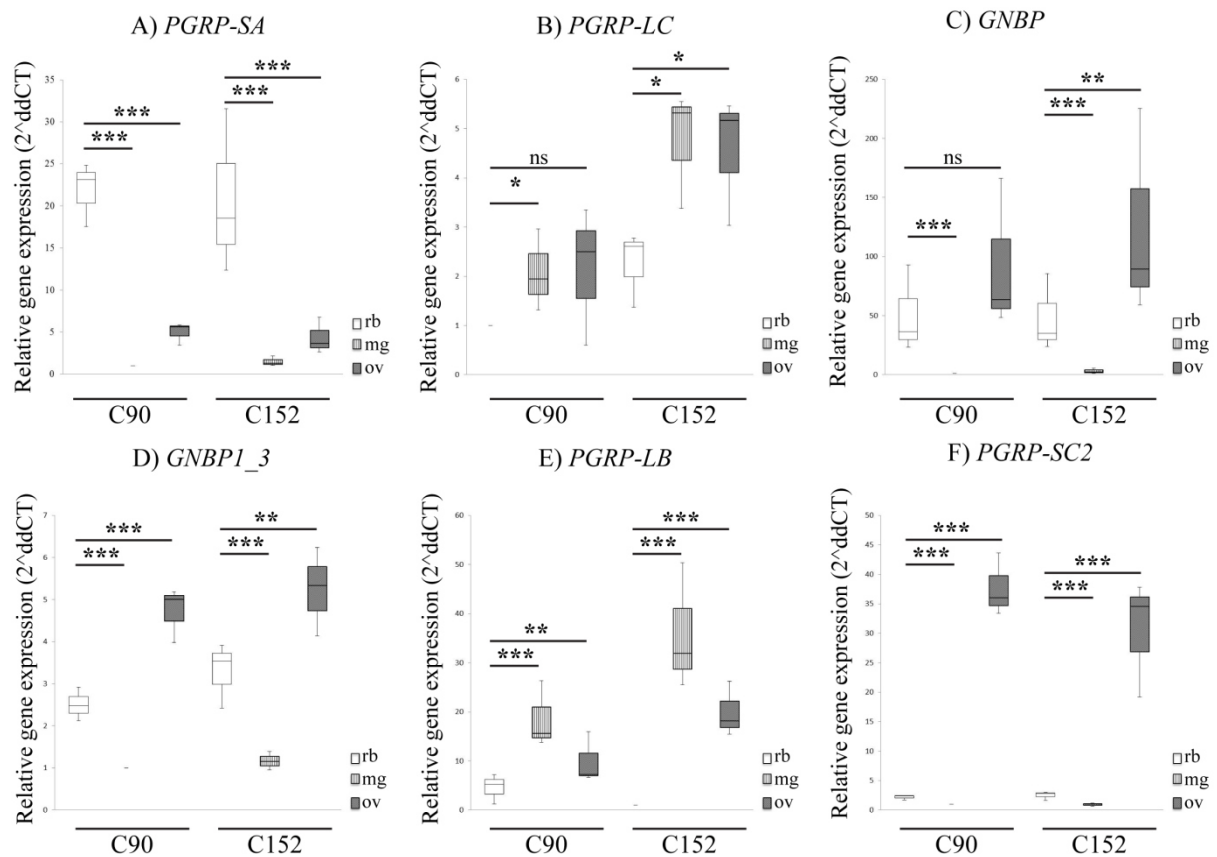
Figure 52 Expression of putative immune genes in different tissues of *C. floridanus*. Ovaries (ov), midguts (mg) and residual bodies (rb) of queenless *C. floridanus* workers were dissected and the expression of the genes **(A)** *peptidoglycan-recognition protein PGRP-SA*, **(B)** *PGRP-LC*, **(C)** *beta-1,3-glucan-binding protein GNBP*, **(D)** *GNBP1_3*, **(E)** *PGRP-LB* and **(F)** *PGRP-SC2* was analysed via qRT-PCR. Average gene expression levels relative to the housekeeping gene *rpL32* (dCt = Ct(target) - Ct(*rpl32*)) were determined and tested for significant differences between the tissues using an ANOVA (Table 27) followed by an LSD post hoc test (Table 28). The Box-plots show normalised changes in gene expression relative to the expression data set with lowest expression level (2^{ddCt}-values). For easier representation of differences in gene expression depending on the colony, the relative gene expression within tissues is presented separately for the two colonies C90 and C152. Significant differences in gene expression levels according to the LSD post hoc test are marked with asterisks (*: p < 0.05; **: p < 0.01; ***: p < 0.001; ns = not significant).

The gene encoding the PGRP-LC (Cflo_N_g10272t1), which is predicted to recognise DAP-type peptidoglycan of Gram-negative bacteria, was about 2-fold higher expressed in midgut and ovarian tissue and showed lowest expression levels in the residual body (C90: p(rb/mg) = 0.0226, p(rb/ov) = 0.0594 (ns); C152: p(rb/mg) = 0.0115, p(rb/ov) = 0.0166) (Figure 52). Additionally, two PRR genes encoding the putative negative immune regulators, PGRP-LB (Cflo_N_g103t1) and PGRP-SC2 (Cflo_N_g102t2), were significantly more strongly expressed in the ovaries than in the residual body (Figure 52). The expression of *PGRP-LB* appeared to be highly dependent on the factor "Tissue" (p < 0.0001) and the interaction "Tissue*Colony" (p = 0.0011). While its expression was only 2-fold higher in the ovaries of colony C90 (p(rb/ov) = 0.0077) and about 3-fold higher in the midgut tissue (C90: p(rb/mg) = 0.0001), *PGRP-LB* was about 20-fold more strongly expressed in the ovaries (p(rb/ov) < 0.0001) and 30-fold more strongly expressed in the midguts (p(rb/mg) < 0.0001) of workers from the colony C152. These results confirm previous data which detected a highly significant about 6-fold induced expression of *PGRP-LB* in the midguts of *C. floridanus* queenright workers W2 (Ratzka, 2012). Interestingly, the expression of *PGRP-SC2* within the tested tissues of queenless workers was only dependent on the effect "Tissue" (p < 0.0001) and maximal expression levels were observed in the ovarian tissue when compared to the residual body (C90: p(rb/ov) < 0.0001; C152: p(rb/ov) < 0.0001). Furthermore, lowest expression levels of *PGRP-SC2* were observed in the midgut tissues (C90: p(rb/mg) = 0.0008; C152: p(rb/mg) < 0.0001) which confirmed previous findings in the midgut tissue of queenright workers, where an about 9-fold lower expression of *PGRP-SC2* was detected (Ratzka, 2012).

The gene *TEP2* (Cflo_N_g4492t1) encodes one of three identified thioester-containing proteins in *C. floridanus*. TEP2 belongs to the putative effectors of the Jak-Stat pathway and supposedly promotes phagocytosis (Blandin and Levashina, 2004; Bou Aoun et al., 2011). Although the expression of *TEP2* in queenless workers of *C. floridanus* was dependent on the effect "Tissue" and the interaction "Tissue*Colony", similar expression patterns were observed for both colonies (Figure 53). *TEP2* was most strongly expressed in the residual body tissue, and significantly lower expression levels were observed in the ovaries of animals from colony C90 (C90: p(rb/ov) = 0.0034; C152: p(rb/ov) = 0.0623, ns). Within the midgut tissue of both colonies, *TEP2* exhibited similarly low expression levels (C90: p(rb/mg) < 0.0001; C152: p(rb/mg) < 0.0001; p(C90mg/C152mg) = 0.9678, ns). Overall, the results indicate a downregulation of *TEP2* expression in the endosymbiont bearing tissues midgut and ovary.

The expression of the gene encoding the serine protease Persephone (Psh; Cflo_N_g8442t1) was significantly upregulated upon infection of *C. floridanus* with Gram-positive and Gram-negative bacteria (chapter 4.1.4). These findings strongly suggest that Psh is an active component of the Toll signalling pathway, which is mainly activated in response to infection with Gram-positive bacteria or fungi. In queenless workers of *C. floridanus*, lowest expression levels of *psh* were observed in the

midguts (C90: p(rb/mg) < 0.0001; C152: p(rb/mg) < 0.0001). Also, the gene was significantly lower expressed in ovaries when compared to the expression within the residual body (C90: p(rb/ov) < 0.0001; C152: p(rb/ov) < 0.0001). *Psh* was also significantly higher expressed in the ovaries when compared with the expression levels in the midgut tissue (C90: p(mg/ov) < 0.0001; C152: p(mg/ov) < 0.0001) which may suggest a low Toll signalling pathway activity in the ovaries of queenless *C. floridanus* workers.
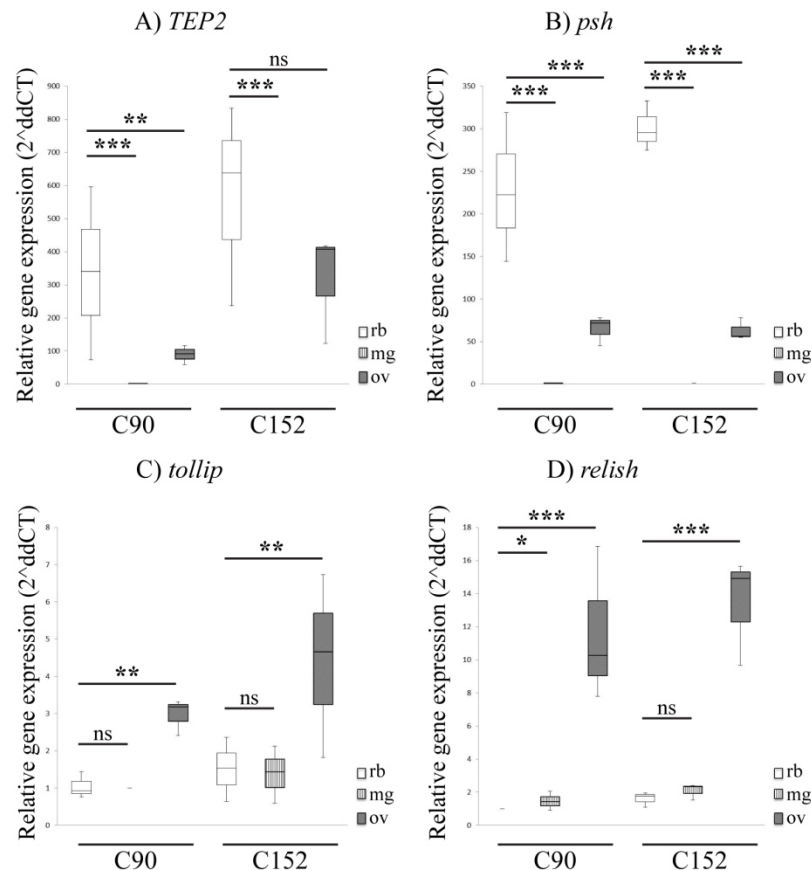
Figure 53 Expression of putative immune genes in different tissues of *C. floridanus*. Ovaries (ov), midguts (mg) and residual bodies (rb) of queenless *C. floridanus* workers were dissected and the expression of the genes **(A)** *alpha-2-macroglobulin-like protein TEP2*, **(B)** *serine protease persephone psh*, **(C)** *Toll-interacting protein tollip* and **(D)** *Nuclear factor NF-kappa-B p110 subunit relish* was analysed via qRT-PCR. Average gene expression levels relative to the housekeeping gene *rpL32* (dCt = Ct(target) - Ct(*rpl32*)) were determined and tested for significant differences between the tissues using an ANOVA (Table 27) followed by an LSD post hoc test (Table 28). The Box-plots show normalised changes in gene expression relative to the expression data set with lowest expression level (2^{ddCt}-values). For easier representation of differences in gene expression depending on the colony, the relative gene expression within tissues is presented separately for the two colonies C90 and C152. Significant differences in gene expression levels according to the LSD post hoc test are marked with asterisks (*: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$; ns = not significant).

In weevils, a homolog of the Toll-interacting protein Tollip was more strongly expressed in the bacteriome, the organ which harbours the weevil endosymbionts, than in the residual body parts. It was suggested that Tollip acts as a negative regulator of the immune response toward the endosymbionts (Anselme et al., 2008). A previous study revealed no significant differences and therefore equal expression levels of *tollip* (Cflo_N_g363t1) in the midgut and residual body of different developmental stages of *C. floridanus* (Ratzka, 2012). These results were confirmed in the

current study. The expression of *tollip* showed no significant differences between the midgut tissue and the residual body (C90: p(rb/mg) = 0.9761; C152: p(rb/mg) = 0.7602). However, *tollip* was significantly and about 3-fold higher expressed in the ovarian tissue when compared to the expression levels in the residual body (C90: p(rb/ov) = 0.0021; C152: p(rb/ov) = 0.0022) which suggests that the putative negative regulator Tollip might be involved in the modulation of the immune response toward tolerance of *B. floridanus* in *C. floridanus* ovaries.

Additionally, highest expression levels of *relish* (Cflo_N_g6082t1) were observed in the ovarian tissue of queenless workers. The gene was significantly and about 10-fold more strongly expressed in the ovaries when compared to expression in residual body parts (C90: p(rb/ov) < 0.0001; C152: p(rb/ov) < 0.0001) and showed only a slightly higher expression in the midgut tissue of the ants (C90: p(rb/mg) = 0.0469; C152: p(rb/mg) = 0.0934, ns). The significant differences in *relish* expression between midgut and ovarian tissue in queenless *C. floridanus* workers (C90: p(mg/ov) < 0.0001; C152: p(mg/ov) < 0.0001) suggest that the transcription factor might play a role in symbiont tolerance in ovaries but not in midguts.

5 Discussion

5.1 The immune gene repertoire of *C. floridanus*

The genome of *C. floridanus* comprises a broad immune repertoire including almost all components of the main immune signalling pathways. However, only a relatively small number of peptidoglycan recognition proteins and known antimicrobial peptides were identified. As it was discussed for social insects in general (Otti et al., 2014), the ants may not only rely on an AMP-based internal immunity, but measures collectively referred to as external and social immunity may play complementary roles for the complete immune repertoire of *C. floridanus*.

In a previous study, 35 infection-inducible genes from *C. floridanus* were identified utilising suppression subtractive hybridisation (SSH). Genes involved in pathogen recognition and signal transduction, and genes encoding effectors with putative antimicrobial activity were upregulated upon injection of Gram-positive and Gram-negative bacteria including the endosymbiont *B. floridanus* (Ratzka et al., 2011). The results of subsequent investigations indicated that the immune response in the ants might be modulated in order to enable tolerance of the endosymbionts, which can still stimulate the host immune system (Ratzka et al., 2013a). However, the previously published SSH study included only a limited number of immune genes. In order to identify further immune-relevant and putative symbiosis-relevant genes, a transcriptome study was performed. For this purpose, samples from immune challenged and untreated larvae and workers were applied for the transcriptome analysis via RNA sequencing. The results obtained were used for a comprehensive re-annotation of the previously published genome of *C. floridanus* (Bonasio et al., 2010). The enhanced quality of the annotation allowed the improvement of gene prediction and their refined functional annotation, and the *C. floridanus* immunome was generated utilising sequence-based protein orthology with *D. melanogaster*, *A. mellifera* and *N. vitripennis* (Gupta et al., 2015). Most importantly, the genes encoding proteins with immune-related function comprise genes involved in the major categories of the immune system, such as microbial recognition and signalling pathways, phagocytosis, encapsulation and melanisation, antiviral defence and AMPs. As described for *N. vitripennis* (Sackton et al., 2013), effector molecules seem to be less conserved and appear to be taxon-specific, whereas most components of the immune signalling pathways were found to be highly conserved.

Orthologs of almost all components known from the Toll, IMD, JNK and Jak-Stat signalling pathways in *D. melanogaster* were identified in the genome of *C. floridanus*. Three pattern recognition proteins (PRRs) which activate the Toll signalling pathway upon recognition of the lysine-type peptidoglycan (PGN) from Gram-positive bacteria or fungal glucans were found: the peptidoglycan recognition protein PGRP-SA and two Gram-negative binding proteins (GNBPs), with one of the GNBPs showing significant upregulation of gene expression upon infection. Additionally, the expression of the gene encoding the serine protease Persephone was significantly upregulated upon bacterial immune

challenge. In *Drosophila*, the Psh-mediated activation of the Toll signalling pathway was shown to be independent from the recognition of cell wall structures. Instead, Psh was shown to recognise bacterial or fungal virulence factors, such as proteases and chitinases of the fungus *Metarhizhium anisopliae* (Gottar et al., 2006), or proteases secreted by Gram-positive bacteria (Ashok, 2009; El Chamy et al., 2008). In *C. floridanus*, the expression of *psh* was significantly upregulated upon infection with a mixture of heat-killed Gram-positive and Gram-negative bacteria. Therefore, it can be assumed that Psh stimulates the activation of Toll signalling. However, the exact bacterial factors which are recognised by the serine protease need to be further investigated. Following the recognition of either lysine-type PGN or several virulence factors, the serine protease cascade of the Toll pathway is initiated and results in the activation of Spätzle upon cleavage by the Spätzle processing enzyme SPE. Spätzle then binds to the membrane-bound receptor protein Toll which initiates an intracellular signalling cascade upon dimerization. Although most of the extracellular and intracellular components of the Toll pathway were identified in the *C. floridanus* genome, only one gene encoding the NF-κB transcription factor Dorsal was found, whereas the genome lacks the gene coding for the transcription factor DIF. Similarly, no ortholog of DIF was found in the *A. mellifera* genome. The absence of a DIF ortholog in the genome of *C. floridanus* suggests that DIF may be a derived protein whose homologs are possibly found only in brachyceran flies (Evans et al., 2006). In *C. floridanus*, Dorsal appears to be the sole transcription factor required for AMP expression upon Toll-activation.

Whereas the *Drosophila* genome encodes two PGRPs recognising the DAP-type PGN of Gram-negative bacteria, only one such PGRP was found in the genome of *C. floridanus*. This PGRP, now annotated as PGRP-LC, shows high homology to the respective proteins in *D. melanogaster* and *A. mellifera*. Hence, it can be assumed that the membrane-integrated PGRP-LC is the only receptor triggering the IMD pathway, since an ortholog of the soluble PGRP-LE of *Drosophila* was not identified. Interestingly, the *Drosophila* PGRP-LE activates the IMD pathway upon recognition of extra- and intracellular PGN (Kaneko et al., 2006; Kurata, 2010). Therefore, it can be argued that *B. floridanus* might be able to escape the detection via the host immune system due to the intracellular localisation of the endosymbionts and the lack of the possibly intracellularly active PGRP-LE. Furthermore, two genes encoding PGRPs with putative regulatory function were identified, *PGRP-SC2* and *PGRP-LB,* with the latter showing significant upregulation after an immune challenge. The two PGRPs comprise amidase domains which probably enable the cleavage of DAP-type PGN into non-immunostimulating molecules and may therefore lead to a down-modulation of the IMD signalling, as it was shown for the *Drosophila* proteins (Bischoff et al., 2006; Kurata, 2014; Royet et al., 2011). An ortholog of the regulatory PGN recognition protein PGRP-LF of *Drosophila* was not found in the *C. floridanus* genome. In the fruit fly, it was shown that PGRP-LF does not recognise PGN but rather binds to the PGRP-LCx ecto-domain, which normally binds and presents PGN to the recognition domain PGRP-LCa. Therefore, it was suggested that PGRP-LF downregulates the IMD signalling pathway on the basis of the competitive interaction with the PGN recognition protein

PGRP-LC (Basbous et al., 2011). However, the two regulatory PGRPs encoded by the *C. floridanus* genome may be sufficient for an effective down-modulation of the IMD pathway mediated immune response. In fact, it was shown that *PGRP-LB* expression levels are highly upregulated only in the midgut tissue of pupae, and it was suggested that the massive multiplication of *B. floridanus* in pupae may be supported by the PGRP-LB mediated down-modulation of the IMD pathway in the midgut. Therefore, the identified amidase-containing PGN recognition proteins may not only prevent an overactive IMD signalling pathway but may also play a role in endosymbiont tolerance in the ants (Ratzka et al., 2013a). The IMD pathway in *C. floridanus* lacks the regulatory Kenny subunit of the IKK complex. This might be a common characteristic of the hymenopteran IKK complex, since the genomes of *A. mellifera*, *N. vitripennis* and other ant species also lack the Kenny subunit. Additionally, Caudal, which is a negative regulator of the transcription factor Relish, was not identified in the *C. floridanus* genome. Caudal was suggested to play a role in maintenance of tissue homeostasis in the *Drosophila* midgut (Ryu et al., 2008).

Almost all components involved in JNK signalling were identified in the genome of *C. floridanus*. The presence of the JNK signalling pathway suggests that the ants are therefore able to induce gene expression relevant for wound healing, for inflammatory and cellular immune responses, as it has been described for *Drosophila* (Rämet et al., 2002a; Sluss et al., 1996). However, it is quite surprising that the *C. floridanus* genome lacks all three known ligands stimulating the Jak-Stat pathway in *Drosophila*. It was shown that the ligands are also absent from the genomes of other insects (Evans et al., 2006). More importantly, the expression of the gene encoding the thioester-containing protein TEP2 containing a conserved CGEQ motif, which was shown to promote covalent binding to the bacterial surface (Bou Aoun et al., 2011), was found to be significantly upregulated upon immune challenge in *C. floridanus* larvae. Although the components stimulating the Jak-Stat signalling pathway in the ant are unknown, these findings suggest that the pathway should be functional and may lead to the infection-induced expression of effectors such as TEPs, which were shown to be involved in phagocytosis in other insects (Blandin et al., 2004; Bou Aoun et al., 2011).

As mentioned above, the repertoire of antimicrobial peptides of insects appears to be highly species-specific. Only a small number of known AMPs was identified in *C. floridanus*, and the ant shares homologs of hymenoptaecin, two defensins, tachystatin-like peptides and two putative crustin-like peptides/ waprins with other ant species as well as with *A. mellifera* and *N. vitripennis*. Especially in comparison to the AMP repertoire of the solitary wasp *N. vitripennis*, the number of AMPs in *C. floridanus* and other social insects appears to be quite low. However, the special structure of the multipeptide precursor of hymenoptaecin should be taken into consideration. The *C. floridanus* hymenoptaecin precursor molecule contains six domains with high similarity to hymenoptaecin as well as a single hymenoptaecin-like domain with an N-terminal extension and few other sequence modifications. It is likely that the mature peptides are released via proteolytic cleavage upon immune

challenge of the ant (Ratzka et al., 2012a). Similar structural features have been described for the hymenoptaecin multipeptide precursors of other ant species (Zhang and Zhu, 2012). Additionally, the expression of the *hymenoptaecin* gene was strongly induced in immune-challenged *C. floridanus* in a previous study (Ratzka et al., 2011) as well as in the validation of the transcriptome study of this work (chapter 4.1.4; (Gupta et al., 2015)). As suggested before (Ratzka et al., 2012a), the immune-related induction of hymenoptaecin expression and the cleavage-mediated release of seven hymenoptaecin peptides with putative antimicrobial function may result in an amplification of the immune response in *C. floridanus*. Consequently, hymenoptaecin genes, which so far have been identified in hymenopterans only, might generally result in a functional extension of the AMP repertoire of the ants or hymenopterans in general.

Additionally, a trade-off between external and internal immune defence has been widely discussed (Otti et al., 2014). External immunity comprises both, antimicrobial secretions and behavioural adaptations. The secretion of antimicrobial substances, such as the antibiotic metapleural gland secretions in the leaf-cutting ants *Atta columbica* and *Acromyrmex octospinosus* (Fernandez-Marin et al., 2006; Poulsen et al., 2002), plays an important role in self and group sanitation and therefore in the protection from the spreading of an infection on colony level. It is interesting to note that the metapleural gland is absent in the vast majority of ant species of the genus *Camponotus* (Hölldobler and Engel-Siegel, 1984; Yek and Mueller, 2011). Also, several other measures, such as salivary gland secretions in termites (Bulmer et al., 2009) and oral secretions of Royal jelly in the honey bee (Romanelli et al., 2011), support the preservation of food and the nests (Otti et al., 2014). Besides chemical defences, behavioural adaptations are of importance to encounter the increasing disease risk with sociality due to higher densities and numbers of individuals which are often close relatives and therefore most likely susceptible to the same parasitic or pathogenic infection (Cremer et al., 2007; Stow and Beattie, 2008). Several studies could demonstrate preventive and therapeutic immune-relevant behaviour in ants, such as nest compartmentalisation and separation of garbage from food sources or fresh brood (Ballari et al., 2007; Hart and Ratnieks, 2001), and the removal of infected brood (Tragust et al., 2013). Recent studies showed increased trophallactic behaviour in immune-challenged carpenter ants of the species *Camponotus pennsylvanicus* and *C. fellah*. This may be a mechanism by which social immunisation and enhanced disease resistance at the colony level are achieved (de Souza et al., 2008; Hamilton et al., 2011). However, the complementary or additive role of such social measures in relation to internal immune defence mechanisms in *C. floridanus* has not yet been analysed in detail and may therefore be of interest for future studies.

Furthermore, the multi-layered defence system comprises another arm, the so-called RNA interference (RNAi), which is involved in gene regulation, the protection of the genome, and anti-viral responses (Hussain and Asgari, 2014). It was shown that modulation of host miRNAs occurs in insects infected with endosymbiotic bacteria. In *Wolbachia*-infected mosquitoes, the miRNA aae-miR-2940 was

highly abundant when compared to uninfected mosquitoes and shown to positively regulate the metalloprotease transcript level (Hussain et al., 2011) and negatively regulate transcription levels of the DNA methyltransferase AaDnmt2 (Zhang et al., 2013). In contrast, overexpression of AaDnmt2 significantly reduced *Wolbachia* density and enhanced virus replication in the mosquitoes. Therefore, it was suggested that the *Wolbachia*-induced and miRNA-mediated downregulation of AaDnmt2 may contribute to inhibition of virus replication and to manifestation of the endosymbionts in their hosts (Zhang et al., 2013). Several miRNAs were differentially expressed upon bacterial infection in *A. mellifera* and predicted to potentially regulate genes involved in immune signalling, melanisation, and AMP expression (Lourenco et al., 2013). In *C. floridanus* genes encoding the core components of the RNAi pathway, such as Dicer or RISC proteins, are well conserved (Bonasio et al. 2010), and it is likely that the pathway is involved in the ants' immune response.

Overall, utilising a transcriptome analysis allowed the identification of a broad immune repertoire in *C. floridanus* showing that most components absolutely fundamental for a comprehensive immune response are present in the genome of the ant. In addition, factors putatively involved in endosymbiont tolerance were identified, e.g. PGRP-LB and PGRP-SC2. However, other branches of the immune system, such as RNA interference and social immunity, have not been analysed in this work. Therefore, future studies of the *C. floridanus* immune system may also focus on these external and internal immune defences in order to understand the regulatory mechanisms to maintain a "chronic" infection with the endosymbiont *B. floridanus* while being able to efficiently combat pathogenic infections.

5.2   The infection-induced differential expression of immune-related genes in *C. floridanus*

257 genes were differentially expressed upon immune challenge in *C. floridanus*. However, most genes showing differential expression after an infection are involved in recognition or encode immune effectors rather than signalling components. For validation of the transcriptome data, the expression of several genes with putative immune function was analysed via qRT-PCR. The comparative analysis of infection-induced gene expression in two different developmental stages of *C. floridanus* did not only confirm the results obtained by the whole transcriptome study based on Illumina sequencing, but additionally suggests a stage-specific induction of immune gene expression in the ants.

After immune challenge with a mixture of heat-killed Gram-negative and Gram-positive bacteria, several genes involved in the *C. floridanus* immune response were upregulated. These genes encode pattern recognition receptors like the Gram-negative binding protein GNBP, several serine proteases, e.g. Stubble and Snake, proteins involved in downstream-signalling and the induction of transcription, for example the NF-κB p110 subunit Relish, as well as immune effectors such as the AMP defensin-1. Additionally, several genes coding for negative immune regulators were found to be upregulated upon infection, e.g. *PGRP-LB* and the gene encoding the NF-κB inhibitor Cactus. Interestingly, most of the genes encoding components of the immune signalling pathways Toll, IMD and Jak-Stat, though identified due to the presence of the respective transcripts in the RNA samples from untreated and immune-challenged ants, were not differentially expressed upon infection. It has been shown before that genes involved in recognition and signalling are less likely to be upregulated upon infection than effector genes (Sackton et al., 2013). In the parasitoid wasp *N. vitripennis*, three out of twelve PGRPs and especially genes encoding for AMPs were significantly upregulated after immune challenge with *Serratia marcescens* and *Enterococcus faecalis*. In contrast, none of the genes encoding signalling proteins showed differential expression (Sackton et al., 2013). Furthermore, the transcriptional response in honey bee larvae revealed that neither the Toll receptor genes nor the gene encoding the universal Toll adaptor, MyD88, were differentially expressed at 24 hours after a fungal infection. However, additional qRT-PCR analysis highlighted that the transcript level of MyD88 showed no changes at 24 hours post injection, but the gene expression was highly upregulated at 36 hours after the infection. This indicates that signalling components of the Toll pathway may show different expression kinetics (Aronstein et al., 2010). For Illumina sequencing and the validation of the immune transcriptome of *C. floridanus*, RNA samples of untreated and immune-challenged animals were taken 12 hours after immune challenge. It can be assumed that most of the genes coding for components of the immune signalling pathways are either constitutively expressed or the transcription activation of these genes may not have been detectable at the chosen time point after the bacterial infection. Additionally, whole animals were used for RNA isolation similarly to a transcriptome study in *A. mellifera*. It was discussed that genes involved in Toll signalling are expected to be upregulated in a subset of tissues rather than in the whole animals (Aronstein et al., 2010). Hence, the regulation of

immune gene expression might have been masked in the total RNA samples, and it may be of future interest to analyse the infection-induced changes in immune gene expression especially in the fat body, which is the main immune organ of insects.

In immune-challenged *C. floridanus*, several genes involved in the general stress response and detoxification were differentially expressed in addition to genes with known and putative immune function. The upregulation of stress-related genes encoding proteins, such as cytochromes, transferrins, the Apolipoprotein D/LIM domain kinase 1 and heat shock proteins, upon immune challenge in the ants has been reported before (Ratzka et al., 2011). The expression of these genes appears to be related to the introduction of septic injuries, as it has been shown for several other insects like the flour beetle *Tribolium castaneum*, the whitefly *Bemisia tabaci*, the burying beetle *Nicrophorus vespilloides* or the pea aphid *Acyrthosiphon pisum* (Altincicek et al., 2008; Gerardo et al., 2010; Vogel et al., 2011; Zhang et al., 2014a). Stress and immune responses seem to be partially linked, as it was demonstrated in *Drosophila*. The immune-related IMD pathway activation also stimulates the JNK pathway, which is associated with stress responses such as cytotoxic stress and apoptosis (Davies et al., 2012; Takeda et al., 2008). The upregulation of stress-related genes and the presence of genes coding for almost all components of the IMD, JNK and Jak-Stat pathways in the *C. floridanus* genome suggest that the infection-mediated activation of the immune response very likely also leads to the expression of genes encoding chaperons, detoxification enzymes, and other stress-related proteins which may help maintaining cellular homeostasis, as has been discussed for *Drosophila* and the whitefly *B. tabaci* (Davies et al., 2012; Zhang et al., 2014a).

In contrast to the upregulation of immune-related genes and of genes related to stress responses, the expression of several genes involved in metabolism and storage was downregulated upon immune challenge, which might indicate a trade-off between different energy intensive processes in *C. floridanus*. Bacterial infections lead to the impaired expression of genes encoding the storage proteins Vitellogenin and Hexamerin 70 in *A. mellifera* (Lourenco et al., 2009; Lourenco et al., 2012) or two putative storage protein genes, *ScSP1* and *ScSP2*, in the silkworm *Samia cynthia ricini* (Meng et al., 2008). As it was suggested for other insects, the downregulation of energy intensive but temporarily dispensable gene expression might support the efficient induction of an immune response in the ants.

Most interestingly, a stage specific induction of immune gene expression was observed in immune-challenged *C. floridanus*. The qRT-PCR analyses were performed to validate the results obtained via Illumina sequencing. Therefore, expression levels of several genes were analysed separately in larvae and adult workers. The vast majority of the tested genes showed similar expression patterns as found in the Illumina-based transcriptome analysis in at least one of the developmental stages. The statistical analyses of the qRT-PCR data revealed the significant influence of both factors, "treatment" (immune challenge or untreated control) and "developmental stage". For example, expression of the genes *NF-*

κB inhibitor cactus, *hymenoptaecin*, *PGRP-LB*, *PGRP-SA* and *thioester-containing protein* (*TEP2*) was strongly induced in immune-challenged larvae, but significantly weaker induced in workers. The stage-specific regulation of gene expression is in agreement with previous studies which showed that the expression of immune-related genes may vary in different developmental stages of holometabolous insects, such as *Bombus terrestris*, *D. melanogaster*, and *A. mellifera* (Colgan et al., 2011; Fellous and Lazzaro, 2011; Randolt et al., 2008).

Furthermore, it was shown that the overall composition of factors involved in the individual immune response may vary between developmental stages. For example, total haemocyte counts were highest in larvae and pupae of the honey bee, while phenoloxidase activity was highest in adults (Wilson-Rich et al., 2008). Also, the immune status of *Formica selysi* care-giving workers significantly influenced the ability of newly eclosed cross-fostered workers to resist a fungal entomopathogen, as highly resistant carers reared more resistant brood when compared to less resistant carers (Purcell and Chapuisat, 2014). In contrast, it was shown that brood care by workers cannot compensate the need of an individual immune competence in *C. pennsylvanicus* larvae (Rosengaus et al., 2013). Larvae which had been vaccinated with heat-killed *S. marcescens* exhibited higher survival rates than naïve animals when challenged with living *S. marcescens*. These effects were independent from brood care by nurse workers (Rosengaus et al., 2013). Taken together, the stage-dependent differential expression of several analysed immune-related genes in *C. floridanus* is in agreement with observations in other insects and might suggest that individuals, although relying on the same set of immune genes, revert to varying combinations of external and internal immunity when facing an acute infection.

5.3 The immune proteome and peptidome of the *C. floridanus* haemolymph

Due to its capability to transport molecules, metabolites, and immune cells throughout the body cavity of the insects, the haemolymph plays an important role in the immune response, in communication, in development, and in physiology. Haemolymph samples of untreated and immune-challenged *C. floridanus* were analysed via mass spectrometry. Besides proteins and peptides involved in the classical immune signalling pathways, several non-canonical proteins with putative immune functions were identified in larval and adult haemolymph. The enrichment of these proteins and peptides upon infection gives further evidence that they might be involved in the immune response of the ants.

1,291 proteins were found in the *C. floridanus* haemolymph samples of which workers and larvae share about 73%. Only about 5% of all identified proteins were specific for worker haemolymph samples, whereas about 22% of the proteins were exclusively detected in the larval haemolymph samples. According to a gene ontology analysis, most of the proteins are involved in metabolism or are predicted to function in the transport of proteins, lipids, and metal ions. About 3% of the proteins are predicted to be involved in immune and stress responses. Furthermore, haemolymph samples were taken from untreated and immune-challenged larvae or adults, since proteins enriched or depleted in the haemolymph of infected insects are very likely to play a role in immunity.

Several extracellularly active components of the classical immune signalling pathways were detected in the haemolymph of *C. floridanus* adults and larvae. These factors comprise pattern recognition receptors (PGRP-SA, PGRP-LB, GNBP and GNBP1_3), several components of the Toll signalling pathway (Snake 2, Persephone, Spätzle), but also putative effectors of the immune response, e.g. lysozymes, AMPs (defensin-1, hymenoptaecin, waprin-Phi1), and the thioester-containing protein TEP2, which might play a role in phagocytosis as shown for other insects (Blandin and Levashina, 2004). Most of these proteins were also enriched upon infection in the haemolymph of either only one developmental stage of the ants or in both, larvae and adults of *C. floridanus*, as it has been shown for homologous proteins in the haemolymph of *D. melanogaster*, *Anopheles gambiae*, and *A. mellifera* (de Morais Guedes et al., 2005; Levy et al., 2004; Randolt et al., 2008; Shi and Paskewitz, 2006; Vierstraete et al., 2004a; Vierstraete et al., 2004b).

Additionally, a non-canonical protein putatively involved in microbial recognition was identified, the Apolipophorin-III. Although the expression level of this protein does not change upon infection, it is one of the most abundant proteins in the haemolymph of *C. floridanus*. Due to its ability to bind microbial cell wall structures, such as lipopolysaccharides (LPS) from Gram-negative bacteria, lipoteichoic acids (LTAs) from Gram-positive bacteria, and fungal beta-1,3-glucan, Apolipophorin-III is considered to be a pattern recognition receptor (PRR) (Halwani and Dunphy, 1997; Halwani et al., 2000; Leon et al., 2006; Ma et al., 2006; Pratt and Weers, 2004; Whitten et al., 2004; Zdybicka-Barabas and Cytryńska, 2013). In *Galleria mellonella*, Apolipophorin-III interacts with LTAs of *Bacillus subtilis*, *Enterococcus hirae* and *Streptococcus pyogenes* (Halwani et al., 2000), but the

protein was also shown to bind to *Bacillus circulans* lacking LTAs, which suggests that the *G. mellonella* Apolipophorin-III can also bind other cell wall components of Gram-positive bacteria (Zdybicka-Barabas and Cytrynska, 2011). It was further demonstrated, that the lipid A and the carbohydrate parts of the *E. coli* LPS are involved in the interaction with the *G. mellonella* Apolipophorin-III (Leon et al., 2006). The binding of Apolipophorin-III to cell wall components leads to their detoxification. Binding to LPS isolated from the outer membrane of insect pathogenic bacteria *Xenorhabdus nematophilus*, reduced the LPS toxicity and prevented *G. mellonella* haemocyte damage (Halwani and Dunphy, 1997). Also, binding of the *G. mellonella* Apolipophorin-III to LTAs from *B. subtilis* was shown to prevent the LTA-induced loss of plasmatocytes, which indicates the protection of the insects against the endotoxin (Halwani et al., 2000). Furthermore, injection of Apolipophorin-III into *G. mellonella* resulted in impaired adhesion of haemocytes but more effective *in vivo* nodule formation (Whitten et al., 2004). The binding of the protein to *Saccharomyces cerevisiae* increased the *in vitro* phagocytic activity of *G. mellonella* haemocytes (Gotz et al., 1997). Although apolipophorin-III is a generally highly abundant protein in the haemolymph of insects, the amount of uncombined apolipophorin-III (apoLpIII) appears to be reduced during stress responses in crickets (Adamo et al., 2008; Weers and Ryan, 2006). Under normal physiological conditions lipids are transported by a high-density lipophorin (HDLp) composed of the two apolipophorins I and II (Weers and Ryan, 2006). However, upon stress-induced release of lipids from the fat body the apoLpIII is recruited and low-density lipophorins (LDLp) are formed to further increase the lipid-carrying capacity of the complex (Weers and Ryan, 2006). As a consequence, the amount of the uncombined apoLpIII with immune function is reduced which results in immunosuppression (Adamo et al., 2008; Weers and Ryan, 2006). A trade-off between lipid transport and immune function of the highly abundant Apolipophorin-III may also be of immunological significance in *C. floridanus* and may be of interest for future research.

Further haemolymph proteins with putative function in the recognition of microbe-associated molecular patterns were identified: Vitellogenins, NPC2-like proteins and the Lectin 4 C-type lectin. Vitellogenins are generally known as storage proteins and yolk precursors, which are synthesised in the fat body and transported into the eggs during oogenesis (Finn, 2007; Zhang et al., 2011). However, it was demonstrated that injection of LPS or LTA into male zebrafish (*Danio rerio*) induced the transcriptional and translational expression levels of Vitellogenin. Additionally, serum Vitellogenin of the zebrafish was able to bind *E. coli* and *S. aureus* thereby inhibiting their growth (Tong et al., 2010). The Vitellogenin from the fish *Hexagrammos otakii* was able to bind to LPS, LTA, and fungal glucans and facilitated lyses of *E. coli* and *S. aureus* as well as phagocytosis of bacteria and fungi by macrophages (Li et al., 2008; Li et al., 2009). Three Vitellogenin proteins were identified in the haemolymph of *C. floridanus*, and Vitellogenin-3 was enriched in adults after bacterial infection, which might suggest a putative role of the protein in immunity.

Two NPC2-like proteins were found to be enriched in the haemolymph of either *C. floridanus* larvae or adults, which might indicate that the proteins play a role in innate immunity, as it has been demonstrated for the homologs in *D. melanogaster*. In the fruit fly, NPC2 proteins bind to LPS, LTAs and PGN (Shi et al., 2012). Additionally, over-expression of NPC2 proteins in PGN-stimulated *Drosophila S*2 cells caused activation of *diptericin* promotors, but not *drosomycin* promotors. While the expression of the AMP diptericin is activated via the IMD pathway, systemic expression of drosomycin is regulated by the Toll pathway. Since the over-expression of NPC2 proteins only increased *diptericin* promotor activity in *S*2 cells stimulated by PGN, it was suggested that these NPC2 proteins may function as PGN co-receptors initiating the IMD signalling pathway, which is required for *diptericin* expression (Shi et al., 2012). Additionally, a calcium-dependent carbohydrate-binding protein, the Lectin 4 C-type lectin, was found in the haemolymph of larvae and adult animals. C-type lectins in insects, for example in the American cockroach *Periplaneta americana* and in *Manduca sexta*, have been shown to be involved in innate immune responses such as phagocytosis and the activation of prophenoloxidase (Jomori and Natori, 1992; Yu and Kanost, 2000; Yu et al., 2002). The synthesis of the best known immulectins from *M. sexta*, IML-2, is highly induced after injection of Gram-negative bacteria or LPS. IML-2 contains two tandem C-type lectin carbohydrate recognition domains, and it was suggested that the two binding sites interact with LPS and initiate the activation of the prophenoloxidase (Yu and Kanost, 2000). Also, inhibition of IML-2 activity resulted in decreased survival rates of *S. marcescens* infected larvae of *M. sexta* (Yu and Kanost, 2000; Yu et al., 2002). The Lectin 4 C-type lectin identified in the haemolymph of *C. floridanus* contains one C-type lectin (CTL)/C-type lectin-like (CTLD) domain including the ligand binding site and was enriched in the haemolymph of immune-challenged adults and larvae, which strongly suggests that the Lectin 4 C-type lectin may be involved in the immune response of *C. floridanus* possibly due to recognition and binding of bacterial structures.

Furthermore, one of the most abundant proteins in the *C. floridanus* haemolymph samples was Hemocytin, which was additionally enriched in the haemolymph of immune-challenged workers. In the silkworm *B. mori*, the gene encoding a homolog to the Hemocytin was expressed during larval-pupal metamorphosis and after injection of *E. coli* or LPS (Kotani et al., 1994), and it was shown that hemocytins in *B. mori* and *Drosophila* are major initiators of nodulation, which means that multiple haemocytes, mostly granulocytes and plasmatocytes, bind to aggregations of bacteria forming an overlapping sheath around the target (Arai et al., 2013; Lavine and Strand, 2002). In *B. mori*, the Hemocytin is accumulated in the granules of granulocytes and only released in response to stimuli such as bacterial infections. The *B. mori* Hemocytin seems to play an important role in the formation of the fibrous nodule matrix (Arai et al., 2013). The high concentration of the homologous protein in haemolymph samples of *C. floridanus* and the infection-induced enrichment of Hemocytin in adults might indicate a role of Hemocytin in either developmental processes or infection-related processes such as nodulation, which connects humoral and cellular components of the immune response.

Discussion

Several studies, including one on *C. floridanus* (Ratzka et al., 2011), show that sterile and septic wounding lead to the upregulation of proteins involved in immune responses as well as in stress responses. Such proteins, e.g. heat shock proteins, cytochromes, glutathione S-transferases or ferritins, play important roles in protein stabilisation, protection against oxidative stresses and insecticides, and may even interact with immune responses, as it was demonstrated for ferritins (de Morais Guedes et al., 2005; Gerardo et al., 2010; Gupta et al., 2015; Ratzka et al., 2011; Shi and Paskewitz, 2006; Vierstraete et al., 2004a; Vierstraete et al., 2004b). Six glutathione S-transferases (GSTs), three heat shock proteins, two transferrins, and three ferritins were found in the *C. floridanus* haemolymph. However, only the Glutathione S-transferase 1-like protein was enriched in the haemolymph of immune-challenged workers, whereas Soma Ferritin was enriched in larval haemolymph after infection. Glutathione S-transferases have been shown to be associated with insecticide resistance (Fang, 2012; Kostaropoulos et al., 2001b; Ramsey et al., 2010; Ranson et al., 2002), while insect ferritins play a role in iron homeostasis and in the protection against oxidative stress (Pham and Winzerling, 2010). The expression of genes encoding ferritins increased after treatment of cultured mosquito cells with iron or hydrogen peroxide and was also induced in animals after a blood meal suggesting a protection from various oxidative challenges (Geiser et al., 2003; Pham and Winzerling, 2010). In addition, it was shown that the expression of ferritin genes is induced in *Tribolium castaneum* upon bacterial infection or injection of crude LPS suggesting a role in innate immunity (Altincicek et al., 2008; Pham and Winzerling, 2010).

In agreement with the results obtained via the immune transcriptome study, several proteins involved in metabolism, e.g. Trypsin-1, Maltase-1, Chymotrypsin-1, Hexamerin, the Zinc metalloprotease ZmpB, and Transferrin, were significantly depleted upon immune challenge. The downregulation of the gene expression and the depletion of proteins involved in storage and metabolism indicates a reduction of these energy-intensive processes upon infection in *C. floridanus*. Similar trade-offs between metabolism and immune responses were observed in other insects, for example *A. mellifera* (Lourenco et al., 2009; Lourenco et al., 2012) and in the silkworm *Samia cynthia ricini* (Meng et al., 2008).

It was already shown that the transcription levels of immune-related genes are influenced by the infection status itself and also depend on the developmental stage of *C. floridanus* (Gupta et al., 2015). This stage-dependent expression of immune genes is reflected in the enrichment and depletion of the respective immune-related proteins in the haemolymph. 645 regulated proteins were identified and found to be enriched or depleted in the haemolymph after infection in either only one of the developmental stages or in both stages. Several proteins, such as the NPC2-like protein (Cflo-N_g7354.t1), the Hemolymph Lipopolysaccharide-binding protein, PGRP-LB, defensin-1, and hymenoptaecin, were enriched in larvae and workers, whereas PGRP-SC2, one GNBP (Cflo_N_g5742.t1) and one NPC2-like protein (Cflo_N_g1858.t1) were enriched in the haemolymph

of workers only. In contrast, PGRP-SA, the TEP2, and Persephone (Cflo_N_g8442.t1) were only enriched in the larval haemolymph. The stage-dependent expression and release of immune-related proteins into the haemolymph have also been observed in the honey bee. The AMPs hymenoptaecin, defensin 1 and abaecin were upregulated in *A. mellifera* larvae and adults in response to bacterial challenge, but only in larvae after they were aseptically wounded. Furthermore, several proteins, for example a phenoloxidase and the PGRP-S2, were only observed in the haemolymph of adult bees when either aseptically wounded or infected with *E. coli* (Randolt et al., 2008).

The current work presents the first study concerning the immune-related haemolymph peptidome of *C. floridanus*. The label-free quantification revealed that a huge number of peptides were enriched or depleted upon infection. Samples from larvae and workers were collected at different time points after immune challenge, and peptides were extracted from the haemolymph for identification via mass spectrometry. For the analysis, only peptides derived from proteins which have putatively functional peptides due to the *in silico* prediction with the SignalP 4.1 Server and the NeuroPred online tool were considered (Nielsen et al., 1997; Southey et al., 2006). Therefore, the analysis in this work focussed on the identification of putative peptides produced via the secretory pathway. Also, peptides which did not show changes in their expression patterns after immune challenge were excluded. In contrast, all peptides with expression changes after immune challenge were categorised into "expression clusters" according to similar expression kinetics based on a semi-quantitative analysis. Previously, it was discussed that an upregulation of antimicrobial peptides late during immune responses is required for killing of bacteria which survived the host's immediate defence mechanisms like phagocytosis or melanisation (Haine et al., 2008). For this work, the ants were immune challenged with heat-killed bacteria. Nevertheless, especially haemolymph peptides which were enriched at later time points after the immune challenge might be immune-relevant.

Taken together, 567 regulated peptides deriving from 39 proteins were identified in the larval haemolymph, and most of the peptides were upregulated at time points later than 24 hours post injection. In the adult haemolymph, 342 regulated peptides deriving from 13 proteins with predicted peptides were identified. In contrast, most of these peptides were upregulated early after an immune challenge and then either depleted or showed continuously high expression at later time points after immune challenge. Most of the identified peptides correspond to either proteins annotated as putative uncharacterised proteins or to the predicted peptides of the AMP hymenoptaecin. However, almost all of the predicted peptides comprise molecular weights > 3 kDa (e.g. around 10-12 kDa for the peptides encoded by the hymenoptaecin-like and hymenoptaecin repeat domains), whereas the identified peptides are generally smaller than 3 kDa. Therefore, most of the identified peptides correspond only partially to the predicted peptides for example of hymenoptaecin or defensin-1. Whether these small peptides are functional or not may be of interest for further research. However, the identified peptide # 9221 (1.26 kDa) in the haemolymph of immune-challenged larvae comprises the complete amino

acid sequence of a predicted soluble peptide of the protein annotated as Vitellogenin and may therefore be a suitable candidate for functional analyses in the future.

For this work, proteins were precipitated in order to separate them from native peptides (< 8 kDa) within the haemolymph samples. Due to the procedure and the identification of peptides with molecular weights only below 3 kDa, it can be assumed that the complete predicted peptides of the known AMPs defensin-1 (about 4 kDa) and hymenoptaecin can rather be found in the protein fractions. Nevertheless, four short peptides derived from defensin-1 were identified in the larval haemolymph, and 484 or 316 short peptides with different amino acid sequences and derived from predicted peptides of the hymenoptaecin (hymenoptaecin-like peptide and hymenoptaecin repeat domain peptides) were found to be regulated in either larval or adult haemolymph. However, the label-free quantification, applied in this work, only revealed differences in the amounts of identified peptides between samples taken at different time points after immune challenge. Hence, it was not possible to distinguish between highly abundant and less abundant peptides derived from the same protein, e.g. hymenoptaecin.

Additionally to hymenoptaecin derived peptides, five peptides which show different amino acid sequences and are derived from the defensin-2 prepropeptide were identified in the haemolymph of adults. These peptides correspond partially to a predicted cleavage product but not to the predicted functional peptide of the defensin-2 (Ratzka et al., 2012a). In accordance, the defensin-2 was not observed in haemolymph protein samples and the expression of the *defensin-2* gene was shown to be downregulated upon infection (Gupta et al., 2015). Thus, the results may indicate a downregulation of the *defensin-2* expression and depletion of the respective protein in the haemolymph of *C. floridanus* upon injection of heat-killed Gram-negative and Gram-positive bacteria.

The analysis of regulated factors upon immune challenge allows the identification of canonical and furthermore non-canonical regulators or effectors of the ant immune system. Although the exact immune-function of a protein or peptide cannot be resolved by its mere abundance in the haemolymph, the current haemolymph proteomic and peptidomic analyses of immune-challenged *C. floridanus* strongly suggest the immune-relevance of several proteins and peptides, and provide candidates for further analyses.

5.4   Expression and functional analysis of hymenoptaecin derived peptides

The *hymenoptaecin* gene of *C. floridanus* encodes a multipeptide precursor protein consisting of a signal peptide and a propeptide which is followed by one hymenoptaecin-like domain and six repeated hymenoptaecin domains. These sequences coding for seven putative hymenoptaecin peptides are separated from each other by putative protease processing sites ("EAEP" and "RR"). Homologs to the yeast KEX proteases are encoded in the *C. floridanus* genome (Bonasio et al., 2010; Gupta et al., 2015), and it was therefore suggested that the hymenoptaecin multipeptide precursor is processed in a similar manner as shown for the yeast α-factor (Ratzka et al., 2012a). The expression of the *hymenoptaecin* gene was upregulated upon infection (chapter 4.1; Gupta et al., 2015; Ratzka et al., 2011), and hymenoptaecin peptides were enriched in the haemolymph of larvae and workers injected with heat-killed Gram-positive and Gram-negative bacteria (chapter 4.2). It was proposed that the immune-induced processing of one multipeptide precursor results in the release of seven putative hymenoptaecin peptides and thus to an amplification of the immune response (Ratzka et al., 2012a). Furthermore, structural differences between the predicted hymenoptaecin-like peptide and the hymenoptaecin repeat domain derived peptides led to the assumption that the hymenoptaecin derived peptides may vary in their function. In this work, two systems for the expression of fusion proteins containing the hymenoptaecin-like peptide (hld) or the hymenoptaecin peptide encoded by one repeat domain (hrep) were established: the baculovirus expression vector system (BEVS) in *Sf*21 insect cells (chapter 4.4.1) and the overexpression of recombinant proteins in an *E. coli* Rosetta 2(DE3) strain (chapter 4.4.3). The Hymenoptaecin derived peptides hld and hrep were purified under native conditions from *E. coli* cell cultures in suitable amounts for initial functional analyses via 384-well plate based antimicrobial activity assays. Whereas hld showed bacteriostatic effects against *E. coli* D31 and *S. aureus*, hrep did not exhibit clear inhibitory effects on the growth of these bacteria.

The major advantage of the BEVS is its ability to produce high yields of authentically processed proteins due to the capacity of insect cells for post-translational processing and modifications, such as glycosylation, phosphorylation and disulfide bond formation (Hericourt et al., 2000; Hodder et al., 1996; James et al., 1995). The MultiBac system is a BEVS specifically designed for eukaryotic multiprotein expression (Berger et al., 2004; Fitzgerald et al., 2006) and was used for the production of recombinant hymenoptaecin derived peptides. Although the respective recombinant baculovirus stocks for protein expression in *Sf*21 cells were successfully generated, only low expression levels were observed for GSThld and 6xHishrep fusion proteins, which is in contrast to the previously described high expression levels of EYFP in Multibac infected *Sf*21 insect cells (Berger et al., 2004). Additionally, the purification with Cytobuster™ reagents revealed the extremely low solubility of the fusion proteins. It was shown before that various recombinant proteins expressed in insect cells exhibit different solubility, although they were expressed under the same conditions (Neuenkirchen, 2012). Also, protein solubility is often impaired due to the shutdown of host protein synthesis and therefore the limited supply of factors such as chaperones which might assist proper folding of recombinant

proteins (Ailor and Betenbaugh, 1999; Hsu and Betenbaugh, 1997). In order to increase solubility and yield of the hymenoptaecin fusion proteins, several other fusion tags, e.g. the S-tag or thioredoxin-tag, could be introduced into the constructs replacing the GST- or 6xHis-fusion tags. It has been demonstrated before that the thioredoxin-tag enhances protein yield and solubility, for example in *E. coli* expression systems (LaVallie et al., 1993; Li, 2011). Additionally, factors supporting proper folding and post-translational modification could be co-expressed to enhance solubility of the recombinant proteins (Hsu and Betenbaugh, 1997).

Due to the low solubility of both fusion proteins, the MultiBac expression system in insect cells was not efficient enough to yield considerable amounts of each hymenoptaecin derived peptide for functional assays. Therefore, the well-established IPTG-inducible pET32-vector expression system in *E. coli* Rosetta strains was used for the expression of thioredoxin (Trx)-S-His6-fusion proteins. The transformation into *E. coli* Rosetta 2(DE3) enabled the codon usage optimisation, since these bacteria additionally carry the plasmid pRARE2, which codes for seven tRNAs recognising rare codons (Kurland and Gallant, 1996; Novy et al., 2001; Sharp and Li, 1986). Although both fusion proteins Trx-S-His6-hld and Trx-S-His6-hrep were soluble and expressed in generally high amounts, expression rates of 40-1,500 mg/litre bacterial cell culture which have been reported for the expression of other proteins (Bogomolovas et al., 2009; Huang et al., 2008; Xu et al., 2007) were not obtained. Furthermore, the purification and digestion of hymenoptaecin fusion proteins yielded into relatively low amounts of 32 µg hld (10.7 µM) out of 4 mg Trx-S-His6-hld hymenoptaecin and about 270 µg hrep (111.4 µM) out of 10 mg Trx-S-His6-hrep. It has been shown before, that the overall recovery of other antimicrobial peptides expressed in *E. coli* decreased with each purification step (Bogomolovas et al., 2009; Huang et al., 2008). Here, it can be assumed that the tag-removal via enterokinase is the crucial step leading to the loss of hymenoptaecin derived peptides. The inefficient cleavage via enterokinase, which also recognises non-canonical cleavage sites, has been observed before (Boulware and Daugherty, 2006; Gasparian et al., 2011). The low recovery of hymenoptaecin derived peptides is very likely the result of inefficient cleavage leading to the production of shorter by-products which might not show bioactivity. To improve the established purification protocol and to obtain a higher overall recovery of purified peptides, alternative cleavage sites recognised by the proteases SUMO or TEV could be introduced into the constructs. Upon tag-removal, both enzymes generate native N-termini without leaving additional amino acids at the N-terminus of the recombinant protein (Kapust et al., 2002; Li, 2011; Waugh, 2011). The removal of the additional amino acids alanine and methionine from the N-terminus of the recombinant hrep might enable amino terminal blocking by the formation of 2-pyrrolidone-5-carboxylic acid out of the N-terminal glutamine residue of native hrep. As it was discussed before, this predicted post-translational modification might further enhance the antibacterial potency of the *C. floridanus* hymenoptaecin peptide hrep (Ratzka et al., 2012a).

Although antimicrobial assays generally require higher concentrations and amounts of antimicrobial peptide stock solutions (Casteels et al., 1993; Casteels et al., 1990; Tonk et al., 2014; Tonk et al., 2015a; Tonk et al., 2015b), the obtained hymenoptaecin peptide (hld/ hrep) stock solutions were suitable for initial growth inhibition assays against the Gram-negative *E. coli* and the Gram-positive *S. aureus*.. Due to the limited amounts of the peptide solutions, *E. coli* and *S. aureus* were only tested once under each condition.

Growth inhibition was observed in *E. coli* D31 samples treated with the recombinant hymenoptaecin-like peptide hld (0.18-1.34 µM) and in *S. aureus* samples treated with 0.01-2.68 µM hld. In contrast, the hrep peptides produced in this work did not exhibit clear inhibitory effects on growth of *E. coli* D31 or *S. aureus*. Minor variations of the measured OD_{600} between untreated and treated *E. coli* D31 samples are likely to be unrelated to treatment, since comparable variations of the OD_{600} were also observed between different wells with untreated control samples. Previous results showed growth inhibition of *E. coli* D31 upon treatment with recombinant *C. floridanus* hrep peptides purified under denaturing conditions (Kupper, 2011; Ratzka et al., 2012a). However, these results could not be confirmed by the current study. Growth of *E. coli* D31 was also inhibited upon treatment with the 1:1 mixture of both peptides (hrep concentrations of 0.87-6.96 µM and hld concentrations of 0.08-0.67 µM). However, inhibition was less than in cultures treated with hld only. Also, growth of *S. aureus* was inhibited when treated with the peptide mixture (hrep $\geq$ 0.05 µM and hld $\geq$ 0.005 µM). Since the treatment with the recombinant hrep peptide alone did not result in an inhibition of bacterial growth, the bacteriostatic effect of the peptide mixture on *S. aureus* and *E. coli* D31 might be more a result of the recombinant hld peptide. Although these preliminary results need to be confirmed in the future, it can be assumed that the recombinant hymenoptaecin-like peptide exhibits bacteriostatic effects on *E. coli* D31 and *S. aureus*. Growth inhibitory effects against various bacteria have already been shown for hymenoptaecin peptides from other insects (Casteels et al., 1993; Gao and Zhu, 2010; Rahnamaeian et al., 2015). Agar inhibition zone assays revealed the broad antibacterial spectrum of the *A. mellifera* hymenoptaecin against several Gram-positive and Gram-negative bacteria including human pathogens and *E. coli*. In liquid culture the hymenoptaecin inhibited bacterial growth at concentrations of 0.5-5 µM (Casteels et al., 1993). The recombinant *N. vitripennis* hymenoptaecin exhibited antibacterial activity against several tested Gram-negative and Gram-positive bacteria depending on the peptides structure. It was shown that the N-terminal fragment of the peptide displayed a significantly weaker antibacterial activity at > 50 µM than the full peptide (2.60 µM) or its C-terminal fragment (2.66 µM) (Gao and Zhu, 2010).

Altogether, the established system for the expression and purification of recombinant *C. floridanus* hymenoptaecin peptides enabled a preliminary functional analysis which indicates that the peptides derived from the hymenoptaecin multipeptide precursor molecule may exhibit activity against different Gram-positive and Gram-negative bacteria. Therefore, the antimicrobial potency of the

predicted hymenoptaecin-like peptide of *C. floridanus* (Ratzka et al., 2012a) was shown. Future analyses of the activity of *C. floridanus* hymenoptaecin derived peptides (hld and various hymenoptaecin repeat domain peptides, which differ in their amino acid sequences) against a broad spectrum of Gram-positive and Gram-negative bacteria as well as fungi will lead to a better characterisation of the antimicrobial repertoire of the ants and may provide potential antimicrobial agents against human pathogens, as it has been shown for the *A. mellifera* hymenoptaecin (Casteels et al., 1993).

5.5 The down-modulation of the immune response in *C. floridanus* ovaries might support vertical transmission of *B. floridanus*

The current work presents a detailed analysis of the distribution of *B. floridanus* in ovarian tissue of *C. floridanus* queenless workers and complements previous studies which focussed mainly on cellular structures of the ovaries, such as accessory nuclei. However, these studies were the first to describe the abundance of bacteria-like symbionts in ants of the genus *Camponotus* (Blochmann, 1882; Buchner, 1918). The bacterial endosymbionts of the ants are vertically transmitted via eggs, as it has been described for other endosymbionts (Balmand et al., 2013; Klein et al., 2016; Kobialka et al., 2015; Matsuura et al., 2012a; Szklarzewicz et al., 2013; Zchori-Fein et al., 1998). The bacteria are transferred into very young determined oocytes via infected follicle cells. Interestingly, and in contrast to endosymbionts in other ants (Klein et al., 2016), *Blochmannia* never reside in the nurse cells. The bacterial endosymbionts could be removed from the ovaries by antibiotic treatment of the workers. Additionally, the expression of several immune-relevant genes in the ovarian tissue of *C. floridanus* was analysed. The results indicate a down-modulation of the immune response in the ovaries, which might support endosymbiont tolerance within this tissue, as it has been shown for the other endosymbiont bearing tissue, the midgut (Ratzka et al., 2013a).

5.5.1 The localisation of *B. floridanus* in ovarian tissue of *C. floridanus*

Six major steps of endosymbiont transmission were identified during oogenesis of the ants: (i) the apical germarium which contains germ-line stem cells (GSCs), cystoblasts and cystocytes is not infected with the bacterial endosymbiont; (ii) in the lower germarium, cystocytes form 16-cell and 32-cell cysts and *B. floridanus* resides within small somatic cells (future follicle cells) lying under the basal lamina, which surrounds each ovariole; (iii) the bacterial endosymbionts rapidly spread from the outer layers into the whole ovariolar tissue infecting follicle cells at first, and *B. floridanus* is transmitted exclusively from follicle cells to the determined oocyte; (iv) the nurse cells are never infected with the endosymbionts; (v) in the lower germarium as well as in the vitellarium, endosymbionts are continuously transmitted into the growing oocytes while the amount of endosymbionts in follicle cells decreases; (vi) finally, *B. floridanus* can only be found in fully grown oocytes, where the bacteria are localised at the posterior pole shortly before the oocyte enters the oviduct (Figure 54). The relatively early infection of determined oocytes ensures the transmission of the endosymbionts into each oocyte/ egg within the polytrophic-meroistic ovarioles and therefore into each embryo of *C. floridanus*. In contrast to the transmission of symbionts in other hymenoptera, nurse cells appear to be never infected with *Blochmannia*, whereas for example *Wolbachia* are concentrated and multiply within the nurse cells of the parasitic wasp *Aphytis*. Through cytoplasmic bridges maternal substances and *Wolbachia* endosymbionts are slowly transmitted into young developing oocytes (Zchori-Fein et al., 1998). A recent study also showed the infection of nurse cells of the invasive ant *Cardiocondyla obscurior* with the endosymbiont *Westeberhardia cardiocondylae,* which

appeared to be transmitted into late-stage oocytes only during nurse cell depletion (Klein et al., 2016). Therefore, the infection of young oocytes via follicle cells in *C. floridanus* broadens the so far observed spectrum of endosymbiont transmission in insects and might be a common characteristic of ants from the tribe Camponotini, as it has also been observed for *Camponotus ligniperda* (Buchner, 1918).
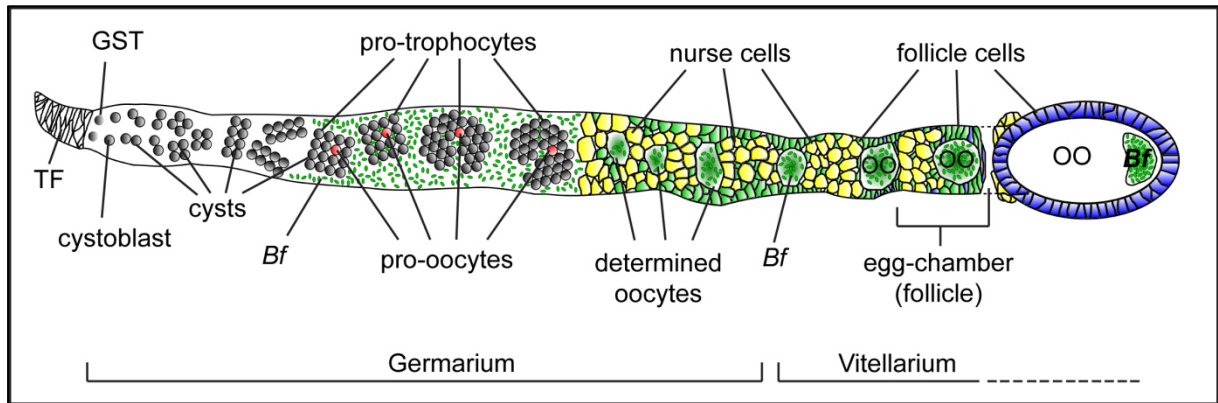
Figure 54 Schematic diagram of the transmission of *B. floridanus* during progressing oogenesis in a *C. floridanus* ovariole. The terminal filament (TF) and the apical germarium, which contains the germ-line stem cells (GSTs), their direct progeny, the cystoblasts and cystocytes forming cysts, are free from *B. floridanus (Bf)*. The "infection zone" is located in the lower parts of the germarium, where the bacteria infect somatic cells lying directly under the basal lamina. *B. floridanus* (green) rapidly infect the whole ovariole's tissue towards posterior, but never pro-trophocytes or pro-oocytes within the cysts. Only with progressing differentiation after the 5th division of cystocytes, one pro-oocyte develops into the determined oocyte (dO) and scattered endosymbionts can be detected in the cytoplasm of dOs, which later on acquire a position in the centre of the ovariolar tube. During egg-chamber formation, nurse cells build packages between dOs, which grow rapidly and develop into oocytes (OO). *B. floridanus* can be detected in the dO and oocytes and in the follicle cells (green) which intersperse nurse cell packages at first and also surround the oocytes. In the vitellarium *B. floridanus* are transported into the growing oocytes, where the bacteria reside within the whole ooplasm at first. In fully grown oocytes, which are surrounded by a single layer of follicle cells (blue) no longer harbouring endosymbionts, Bf can only be observed within the cytoplasm at the posterior pole.

Several observations regarding the cell-to-cell transfer of *B. floridanus* during oogenesis suggest that the endosymbionts "hijack" host exocytosis-endocytosis mechanisms between follicle cells and oocytes in order to be transported into the latter. Within follicle cells and oocytes, the endosymbionts were observed either residing free in the cytoplasm or inside of vacuole-like structures. Similar observations have been made in a previous study describing the localisation of *Blochmannia* within the midgut tissue and in eggs (Sauer et al., 2002). Within the ovariolar tissue, vacuoles were already observed in determined oocytes and cells of somatic origin. The number of vacuoles increased strongly during oogenesis and the highest number of vacuoles was detected in young oocytes when endosymbionts appeared to be transported into the oocytes from the surrounding follicle cells. In fully grown oocytes, only a minority of the endosymbionts was found within vacuoles, and the majority of *B. floridanus* was observed residing freely in the ooplasm. Interestingly, two types of bacteria-containing vacuoles were detected. While most of the vacuoles were surrounded by a single membrane, several vacuoles comprised a double membrane. Additionally, young growing oocytes exhibited numerous widenings of the intercellular space, which most often appeared as invagination-

like structures at the cytoplasmic membrane between follicle cells and the oocyte. Although these observations only display snapshots of endosymbiont transmission and *Blochmannia* was never actually "caught in the act" of transport, it is likely that the described structures play a role in endosymbiont transfer into the oocytes. Some pathogenic bacteria can exploit endocytic pathways in order to access cells thereby avoiding degradation via the endosome-to-lysosome pathway (Asrat et al., 2014; Doherty and McMahon, 2009). *B. floridanus* is not endowed with cell invasion factors known from pathogens (Gil et al., 2003), and it is very likely that the bacterial endosymbionts are transferred via host endocytosis-exocytosis processes from follicle cells into the oocytes. As it has been shown in *Apis mellifera*, Vitellogenin is exocytosed into the narrow perivitelline space between the follicular epithelium and the oocyte and taken up by the oocyte via endocytosis (Fleig, 1995). It can be assumed that similar events take place at the follicle cell-oocyte interface in *C. floridanus* ovaries and therefore may enable the transport of *B. floridanus* into the oocytes. Two possible mechanisms of transport will be discussed in the following.

In the pea aphid *Acyrthosiphon pisum*, *Buchnera* endosymbionts are non-selectively transported into the developing embryo. The endosymbionts are exocytosed from maternal bacteriocytes and thereby temporarily released into the extracellular space. Subsequently, *Buchnera* endosymbionts are internalised by the endocytic activity of the blastula (Koga et al., 2012). The endosymbionts of the cockroach *Blattella germanica* are released from the bacteriocytes in nymphal instars, cross the ovariole sheath and then reside in the extracellular space between the follicle epithelium and the oocyte. They remain in close contact with the oocyte microvilli and are actively phagocytosed by the oocytes shortly before chorion formation (Sacchi et al., 1988). In the hemipterans *Nysius ericae* and *Nithecus jacobaeae*, bacterial endosymbionts are released by the bacteriocytes which closely adhere to the oocytes. On the membrane surface of the oocyte, the endosymbionts penetrate the ooplasm and are enclosed by the oocyte's plasma membrane. Therefore, the bacterial endosymbionts of these hemipterans reside within the ooplasm surrounded by a host-derived membrane, the so-called perisymbiotic membrane (Swiatoniowska et al., 2013). Assuming a similar mechanism of transfer including an extracellular phase of the endosymbionts as described for *Buchnera* and the endosymbionts of *B. germanica* or *Nysius ericae* and *Nithecus jacobaeae*, *B. floridanus* may migrate to the follicle cell-oocyte interface to be exocytosed by the follicle cell into the narrow extracellular space. The endosymbionts would then be taken up by the oocyte via endocytosis and would therefore be enclosed by a single membrane derived from the plasma membrane of the oocyte. However, so far, *Blochmannia* have never been detected in the extracellular space between cells, neither in the midgut nor in the ovaries of the ants.

Other transfer strategies involve budding-like mechanisms as described for the pathogen *Orientia tsutsugamushi* in mites or an actin-based protrusion mechanism as used by the human pathogens *Listeria monocytogenes* and *Shigella flexneri* (Hybiske and Stephens, 2008). According to the

budding-like mechanism, *B. floridanus* would migrate to the plasma membrane of the follicle cell, where the bacteria would be slowly extruded within small buds. Such buds, which are surrounded by a membrane derived from the follicle cell plasma membrane, would be taken up by endocytosis processes of the oocytes. As a consequence, *B. floridanus* would reside within vacuoles with double membranes, and the outermost layer of the vacuole would be derived from the plasma membrane of the oocyte. Alternatively, *B. floridanus* could use the force generated by actin polymerization to protrude from the follicle cell membrane and to force engulfment into the neighbouring oocyte resulting in a bacteria-containing double-membrane vacuoles (Hybiske and Stephens, 2008). Degradation of the vacuole membranes via a state of single-membrane vacuoles would eventually lead to the release of *B. floridanus* into the ooplasm. Also, manipulation of phagocytosis by bacterial factors may prevent the formation of phagolysosomes and the degradation of the endosymbionts. Hence, *B. floridanus* could "escape" into the cytoplasm as it has been demonstrated for several pathogenic bacteria (Asrat et al., 2014; Bayles et al., 1998; Clemens et al., 2009; Haas, 2007; Hybiske and Stephens, 2008).

However, it is possible that the abundance of *Blochmannia*-containing double-membrane vacuoles may not be related to bacterial entry, but instead indicates that autophagocytotic processes (Romanelli et al., 2014) may occur in the oocytes. In *Drosophila*, it was shown that autophagy is not only a cellular pathway for protein and organelle turnover via the lysosomal system but also acts as a part of the innate immune response against pathogens (Yano et al., 2008). In the pea aphid *Acyrthosiphon pisum*, genes encoding i-type lysozyme and cathepsin L (cath L) were highly expressed in the bacteriome tissue, and gene expression was further induced in bacteriocytes of post-reproductive aphids (Nakabachi et al., 2005; Nishikori et al., 2009b). After final ecdysis in winged morph aphids, expression of aphid CPVL, which was localised around *Buchnera* cells in the bacteriocytes, was observed (Nishikori et al., 2009a). It was suggested that CPLV, i-type lysozyme and cath L are involved in the degradation of *Buchnera* in winged aphids via the lysosomal system (Nishikori et al., 2009a; Nishikori et al., 2009b). In cereal weevils (*Sitophilus* sp.), endosymbionts massively multiply in young adults and provide amino acids to their hosts for the synthesis of DOPA, which is required for cuticle formation (Andersen, 2007; Vigneron et al., 2014). Once the protective exoskeleton is achieved, the endosymbionts are rapidly eliminated through activation of host autophagy and apoptosis in a tissue-dependent manner indicating that these cellular processes are involved in the regulation of the endosymbiont population in accordance with host development (Vigneron et al., 2014). In previous studies, *Blochmannia* was occasionally detected in midgut cells other than the bacteriocytes, and these cells harboured higher numbers of SYTO-stained vesicles which might be derived from the lysosomal system (Sauer et al., 2002; Stoll et al., 2010). Further, lysosome-related genes, such as *i-type lysozyme* and *carboxypeptidase vitellogenic-like CPVL*, were highly expressed in the midgut of *C. floridanus* late pupae, whereas the gene encoding cathepsin L was strongly expressed in the midgut of workers. It was suggested that the lysosomal system might be involved in controlling

intracellular symbionts in the midguts (Ratzka et al., 2012b). Therefore, the analysis of autophagy and lysosome-related gene expression in the ovarian tissue will be of interest, as these genes may be involved in control and tolerance of *Blochmannia* (Ratzka, 2012; Ratzka et al., 2012b).

The extensive microscopic analysis carried out in this work provides detailed insights into the bacterial localisation in the ovarioles with respect to oogenesis. However, when and how the endosymbionts are transferred into the ovarian tissue was not yet revealed. Hecht (1923) focussed on the embryogenesis of *C. ligniperdus* and discovered the transfer of the bacterial endosymbionts into cylindrical cells at the posterior pole of the egg in the course of cleavage division. During the formation of the blastoderm syncytium, these so-called primary bacteriocytes were observed to migrate towards the egg's dorsal site and to spread laterally. Together with endodermal cells, primary bacteriocytes would eventually form the midgut. Additionally, endosymbionts were observed in large cells at the posterior egg pole. These cells fused building the blastoderm syncytium and were enclosed by the embryonal band. During segmentation processes, these endosymbiont harbouring cells were located between the midgut and posterior pole of the embryo, in the segment where ovary formation would take place (Hecht, 1923). These observations suggest that the ovarian tissue of *C. ligniperda* becomes infected with bacterial endosymbionts early during embryogenesis. Therefore, the analyses of the localisation and transfer of *B. floridanus* into ovaries of *C. floridanus* during host embryogenesis will be of interest for future studies.

5.5.2   The expression of immune-related genes in *C. floridanus* ovaries

In a previous study, it has been shown that a down-modulation of the immune system in the midgut may occur during pupation, and it was suggested that a reduced immune response supports the massive increase of *Blochmannia* numbers during metamorphosis of the ants (Ratzka et al., 2013a). Therefore, expression levels of several immune-related genes in ovarian tissue of *C. floridanus* were analysed in comparison to expression levels of the respective genes in the midgut tissue, which also harbours the endosymbionts, and the residual body as *Blochmannia*-free reference tissue. In order to avoid constitutive expression of immune-related genes, residual body parts did not contain the fat body. The genes analysed in this work comprised factors involved in recognition and signalling aspects of the immune response: genes encoding four PGN recognition proteins PGRP-SA (Cflo_N_g8526t1), PGRP-LC (Cflo_N_g10272t1), PGRP-LB (Cflo_N_g103t1), PGRP-SC2 (Cflo_N_g102t2), two genes coding for beta-1,3-glucan-binding proteins GNBP (Cflo_N_g5742t1), GNBP1_3 (Cflo_N_g15215t1), and genes coding for alpha-2-Macroglobulin-like protein TEP2 (Cflo_N_g7345t1), the serine protease Persephone (Cflo_N_g8442t1), the Toll-interacting protein Tollip (Cflo_N_g363t1), and the Nuclear factor NF-kappa-B p110 subunit Relish (Cflo_N_g6082t1). Interestingly, the relative expression levels of all genes differed significantly between the three analysed tissues, but expression of some genes was also dependent on the ant colonies from which the samples originated. However, the colony-dependent expression of immune genes reminds of a

previous publication which showed that the expression of antimicrobial peptide genes in the bumblebee *B. terrestris* was significantly dependent on the host line-parasite pairing. Various host lines showed different gene expression patterns upon infection with several isolates of the parasite *Crithidia bombi* (Riddell et al., 2009).

Whereas the expression of *PGRP-SA* was significantly lower in midgut and ovarian tissue of the adult workers, expression levels of the genes encoding GNBP and GNBP1_3 were highest in ovaries and lowest expression was observed in the midgut tissue. PGRP-SA and the two GNBPs might be involved in the recognition of lysine-type PGN from Gram-positive bacteria or fungal glucans leading to the activation of the Toll pathway (Gobert et al., 2003; Gupta et al., 2015; Michel et al., 2001). Therefore, it is unlikely that the upregulation of these PRR genes is correlated with the Gram-negative *Blochmannia* residing in the ovaries or in the midgut tissue, as it was discussed before (Ratzka et al., 2013a). Nevertheless, the differential expression of genes involved in recognition of microbial molecular patterns indicates a general immune activity of the ovarian tissue, which is further supported by the about 2-fold higher expression of *PGRP-LC* in the ovarian and in the midgut tissue when compared to the residual body. In *Drosophila*, the homolog to this PGN recognition protein activates the IMD pathway upon recognition of the DAP-type peptidoglycan of Gram-negative bacteria (Kaneko et al., 2006; Myllymaki et al., 2014). In a recent study, the expression of *PGRP-LC* in midgut tissue showed no differences over developmental stages (larvae, pupae and adult workers). The expression levels in midgut and residual body were almost similar, and only a small but significant increase in *PGRP-LC* expression from larvae to pupae was observed (Ratzka et al., 2013a). The increased expression of *PGRP-LC* in worker ovaries and midguts might indicate the detection of the Gram-negative *B. floridanus*. Though the *C. floridanus* PGRP-LC shows homology to the *Drosophila* PGRP-LC and PGRP-LE, which recognises intracellular bacteria (Kurata, 2014), it is not yet known whether the ant PGRP-LC can recognise extracellular and intracellular bacteria (Gupta et al., 2015).

The expression of the two PRR genes coding for the putative negative immune regulators PGRP-LB and PGRP-SC2 was significantly upregulated in ovarian tissue of the workers. As mentioned before, the protein sequences of both, PGRP-LB and PGRP-SC2, contain an amidase domain including the highly conserved cysteine residue, which is essential for amidase activity. Therefore, both PRRs are homologs to the respective *Drosophila* proteins, which prevent excessive activity of the immune response by downregulation of the IMD pathway via cleavage of DAP-type peptidoglycan (Kurata, 2014; Steiner, 2004). Recently, it was shown that *PGRP-LB* and *PGRP-SC2* were highly expressed in the midgut tissue of *C. floridanus* late pupae, which correlates with the multiplication of *Blochmannia* observed in the midguts during host metamorphosis (Ratzka et al., 2013a; Stoll et al., 2010). In tsetse flies, the endosymbiont *Wigglesworthia* is transmitted via milk gland secretions. Interestingly, the tsetse PGRP-LB is a highly abundant milk protein, and it was suggested that the amidase activity of

PGRP-LB scavenges symbiotic PGN, thus preventing the stimulation of the tsetse immune response and allowing transmission of the endosymbionts (Wang and Aksoy, 2012). In the weevil *Sitophilus zeamais*, wPGRP, an ortholog of the *Drosophila* PGRP-LB, was shown to downregulate the IMD pathway, and expression levels of wPGRP in weevils challenged with Gram-negative bacteria depended on bacterial growth. However, wPGRP expression was generally high in the bacteriome of larvae and upregulated in the symbiotic nymphal phase. During this state, *Sodalis pierantonius* str. SZPE, the primary endosymbiont of these weevils, is released from the host bacteriocytes. Accordingly, it was proposed that wPGRP expression is an adaptive immune response to symbiont release in order to prevent systemic bacterial invasion into insect tissues but also to prevent clearance from the symbionts (Anselme et al., 2006). Hence, it can be assumed that degradation of DAP-type PGN by PGRP-LB and PGRP-SC2 may prevent stimulation of the IMD-mediated immune response in the *Blochmannia*-harbouring tissues including the ovaries, as it has been shown for other endosymbionts and bacterial commensals (Anselme et al., 2006; Ratzka et al., 2013a; Wang and Aksoy, 2012).

Additionally to the differential expression of PRR genes, expression levels of three genes coding for proteins involved in immune signalling were found to be regulated in a way which may support symbiont tolerance within *C. floridanus* ovaries. The expression of the gene coding for the serine protease Persephone (Psh) was significantly downregulated in ovarian and midgut tissue in comparison to the residual body. The *C. floridanus psh* gene shows homology to the *Drosophila* gene *persephone* (GenBank Acc. No.: Dmel\CG6367). In the fruit fly, infection with entomopathogenic moulds triggers a serine protease cascade involving the protease Persephone. The signalling cascade results in the proteolytic activation of the protein Spätzle and therefore in activation of the Toll signalling pathway (Brennan and Anderson, 2004; Ligoxygakis et al., 2002). Interestingly, it was shown that the Persephone-mediated activation of Toll depends on the recognition of virulence factors, such as proteases and chitinases released by fungi and bacteria, rather than on the recognition of cell wall structures (El Chamy et al., 2008; Gottar et al., 2006). In addition to the already mentioned low expression of *PGRP-SA*, the low expression levels of *psh* in the ovarian tissue of adult workers may be further evidence for a down-modulation of Toll signalling in *C. floridanus* ovaries. Furthermore, *tollip* showed equal expression levels in residual body parts and midgut tissue, which is in accordance to previously published results (Ratzka et al., 2013a), but its expression was significantly higher in ovarian tissue. Tollip exhibits significant homologies to a negative regulator of the Toll-like receptor-mediated immune activation in mammals (Zhang and Ghosh, 2002). In the *S. zeamais* bacteriome, *tollip* was shown to be constitutively overexpressed together with another putative negative immune regulator, *wPGRP-1*, as mentioned above (Anselme et al., 2008).

The expression levels of the gene encoding the nuclear factor NF-kappa-B p110 subunit Relish were highest in the ovarian tissue of *C. floridanus* workers but almost similar in the midgut tissue and the

residual body parts. Upon activation of the IMD pathway Relish is translocated to the nucleus and induces the transcription of PGRPs, of negative regulators of the immune response, and of effector genes, e.g. AMP genes (Ganesan et al., 2011; Lemaitre and Hoffmann, 2007). In *Drosophila*, transcription of the gene encoding the PIMS protein, a PGRP-LC interacting inhibitor of Imd signalling, is also activated by Relish. PIMS has been shown to be a negative feedback regulator of the IMD pathway via interaction with the *Drosophila* PRRs PGRP-LC and PGRP-LE, thereby causing their depletion from the plasma membrane and reduced activation of IMD signalling. Therefore, transcription activation via Relish was suggested to prevent hyperactive AMP production upon pathogenic infection and to further support maintenance of gut homeostasis in *Drosophila* by preventing an immune response against commensal bacteria (Ganesan et al., 2011; Lhocine et al., 2008). During metamorphosis of *C. floridanus*, equal expression levels of *relish* were observed in midgut tissue and residual body parts of pupae and adult workers. Only in larvae, expression of *relish* was higher in the midgut tissue than in residual body parts. Interestingly, immune challenge led to an about 10-fold increased *relish* expression in the residual body parts, whereas its expression was only slightly induced in the midgut tissue (Ratzka et al., 2013a). High expression levels of *relish* in the ovarian tissue indicate activity of the IMD pathway which might be stimulated upon recognition of *B. floridanus*. However, whether Relish activates the transcription of negative regulators of the IMD pathway or effectors needs further investigation.

Finally, the low expression levels of the *alpha-2-macroglobulin-like protein TEP2* gene encoding a putative factor belonging to the complement system may contribute to tolerance and transovarial transmission of *B. floridanus*. Lowest *TEP2* expression levels were observed in the midgut tissue, but its expression was also relatively low in the ovarian tissue when compared to residual body parts. The expression of *TEP2* was significantly upregulated in *C. floridanus* larvae upon bacterial infection (Gupta et al., 2015) and the TEP2 protein may be involved in phagocytosis of Gram-positive and Gram-negative bacteria (Bou Aoun et al., 2011; Gupta et al., 2015; Levashina et al., 2001). Therefore, the low expression of *TEP2* in the ovaries may result in reduced activity of the complement system and thereby may reduce phagocytosis of *B. floridanus*.

Taken together, this study suggests a downregulation of Toll signalling via low expression levels of *PGRP-SA* and *psh*, and high expression of the negative regulator *tollip* in *C. floridanus* ovarian tissue. Furthermore, the high expression levels of two PGN recognition proteins (PGRP-LB and -SC2) with regulatory function might reduce IMD pathway activity in ovaries, and high expression levels of *relish* may additionally lead to the transcription of negative regulators of IMD signalling. Additionally, low expression of *TEP2* may reduce the activation of the complement system. These data closely resemble previous observations on immune gene expression in the midgut tissue during metamorphosis of the ants (Ratzka et al., 2013a). Similar to the situation in the midgut, the expression of several immune-

related genes in the ant ovaries appears to be modulated in a way which may support tolerance and transovarial transmission of *B. floridanus.*

5.6   Conclusions and outlook

Within the scope of this thesis a comprehensive overview of the immune gene repertoire of *C. floridanus* and of changes on the transcriptional as well as on protein level in response to immune challenges was obtained. Almost all genes involved in the major immune signalling pathways are present in the *C. floridanus* genome but a comparatively low number of pattern recognition receptors and of known AMPs (Gupta et al., 2015). The low number of AMPs identified here confirms previous observations in *C. floridanus* and other ants (Zhang and Zhu, 2012) but may be compensated by the infection-induced expression of the *hymenoptaecin* gene, which encodes a multipeptide precursor protein. It was predicted that seven peptides with putative antimicrobial activity could be released upon immune-induced proteolytic processing which would lead to an amplification of the immune response (Ratzka et al., 2012a). In this work, two of the predicted hymenoptaecin derived peptides were heterologously expressed and purified. Preliminary functional analyses of both peptides revealed varying activity against the tested Gram-positive *S. aureus* and Gram-negative *E. coli* D31. Therefore, it can be assumed that peptides encoded by the hymenoptaecin precursor molecule exhibit different bioactivity and further extend the spectrum and potency of the ants' immune response. Additionally, a relatively low number of AMPs in comparison to solitary insects seems to be a common feature of social insects such as ants and the honey bee. Future research may also focus on hygiene and other measures referred to as external or social immunity, which may add to and trade-off with the internal immune repertoire of *C. floridanus*, as it was discussed for other social insects (Otti et al., 2014).

The first detailed analysis of the haemolymph proteins of immune-challenged ants largely confirmed the results obtained from the immune transcriptome study. Various canonical proteins, for example components of the classical immune signalling pathways (PGRPs, GNBPs, and serine proteases), but also putative effectors of the immune response (lysozymes and the AMPs defensin-1, hymenoptaecin, waprin-Phi1), and several proteins involved in stress responses were enriched upon infection in the haemolymph. Additionally, non-canonical proteins with putative immune function were enriched upon infection which strongly suggests the immune-relevance of these proteins. Apolipophorin-III, Vitellogenin-3, NPC2-like proteins and the Lectin 4 C-type lectin were enriched in the haemolymph of immune-challenged *C. floridanus*. These proteins may be involved in pattern recognition as it was shown for insects and other animals (Shi et al., 2012; Tong et al., 2010; Yu and Kanost, 2000; Zdybicka-Barabas and Cytryńska, 2013). Therefore, these proteins may extend the range of the comparatively low number of classical PRRs such as PGRPs and GNBPs. In addition, the expression of haemolymph peptides was analysed, and a vast set of peptide candidates was obtained for studies to evaluate their role in immunity. Further analyses of these candidates may help to characterise the complete antimicrobial repertoire of *Camponotus* in the future.

This thesis also represents the first extensive analysis of the localisation of the bacterial endosymbionts in ovarian tissue of *C. floridanus* workers utilising FISH and TEM. *B. floridanus*

initially infects somatic cells in the lower germarium of each ovariole, and the endosymbionts are transferred exclusively from follicle cells into the oocytes during oogenesis. Possible transport mechanisms of the bacteria were discussed, and the abundance of *Blochmannia* residing in vacuoles suggests that the endosymbionts are transferred via entering exocytosis-endocytosis processes between follicle cells and oocytes. An initial trial experiment for preparation of ovarian samples in a close to native state with high pressure freezing could not give further insights into the cell-cell transport of *Blochmannia*, since structural preservation was not improved over the classical fixation protocols. However, this work provides a broad knowledge about the bacterial localisation in the ovarioles with respect to oogenesis and will support the establishment of appropriate conditions for high pressure freezing protocols of ovariole segments of interest.

In the current work, the expression of immune-related genes in ovaries of an ant species was analysed for the first time. Low expression levels of several genes involved in IMD and Toll signalling as well as low expression of a putative opsonin were detected in the ovarian tissue of *C. floridanus*. Interestingly, genes encoding the two amidase PGRPs, PGRP-LB and PGRP-SC2, and a negative regulator of the Toll signalling pathway were found to be highly expressed in the ovaries of queenless workers. Similar to the situation in the midgut (Ratzka et al., 2013a), these new results indicate a down-modulation of the immune response in the ovarian tissue. This may contribute to endosymbiont tolerance in the endosymbiont bearing ovaries and may therefore be important for the successful vertical transmission of *Blochmannia*. To prove this hypothesis, gene knockdown experiments of these negative immune regulators in the ovarian tissue should be performed in the future. Such studies may include molecular methods, such as RNA interference and immunohistochemistry, to further investigate the involvement of immune- and lysosome-related genes in the regulation of endosymbiont populations in the ants, but also to investigate transport mechanisms of the endosymbionts. In addition, the differential infection status of the three major cell types in ovarioles suggests a cell-type dependent expression of immune- and symbiosis-relevant genes. Therefore, the analysis of immune gene expression in ovaries presented in this work provides promising candidate genes for further investigations of cellular gene expression levels via antibody-mediated detection and a previously established protocol for gene knockdown in *C. floridanus* workers (Ratzka et al., 2013b).

Overall, the present thesis contributes to a better understanding of the ants' immune system and its role in establishing a chronic infection with symbiotic bacteria. Improved knowledge about interactions between insects and microbes may eventually provide insights and tools for the development of symbiont-based strategies for the control of insect pests and of insect-derived diseases, as well as for the protection of beneficial insects.

References

Adamo, S.A., Roberts, J.L., Easy, R.H., Ross, N.W., 2008. Competition between immune function and lipid transport for the protein apolipophorin III leads to stress-induced immunosuppression in crickets. The Journal of experimental biology 211, 531-538.

Agaisse, H., Perrimon, N., 2004. The roles of JAK/STAT signaling in *Drosophila* immune responses. Immunological reviews 198, 72-82.

Agaisse, H., Petersen, U.M., Boutros, M., Mathey-Prevot, B., Perrimon, N., 2003. Signaling role of hemocytes in *Drosophila* JAK/STAT-dependent response to septic injury. Developmental cell 5, 441-450.

Aggarwal, K., Silverman, N., 2008. Positive and negative regulation of the *Drosophila* immune response. BMB reports 41, 267-277.

Ailor, E., Betenbaugh, M.J., 1999. Modifying secretion and post-translational processing in insect cells. Current opinion in biotechnology 10, 142-145.

Akman, L., Yamashita, A., Watanabe, H., Oshima, K., Shiba, T., Hattori, M., Aksoy, S., 2002. Genome sequence of the endocellular obligate symbiont of tsetse flies, *Wigglesworthia glossinidia*. Nature genetics 32, 402-407.

Altincicek, B., Knorr, E., Vilcinskas, A., 2008. Beetle immunity: Identification of immune-inducible genes from the model insect *Tribolium castaneum*. Developmental and comparative immunology 32, 585-595.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic acids research 25, 3389-3402.

Anbutsu, H., Fukatsu, T., 2010. Evasion, suppression and tolerance of *Drosophila* innate immunity by a male-killing *Spiroplasma* endosymbiont. Insect molecular biology 19, 481-488.

Anders, S., Huber, W., 2010. Differential expression analysis for sequence count data. Genome biology 11, R106.

Andersen, S.O., 2007. Involvement of tyrosine residues, N-terminal amino acids, and beta-alanine in insect cuticular sclerotization. Insect biochemistry and molecular biology 37, 969-974.

Anselme, C., Perez-Brocal, V., Vallier, A., Vincent-Monegat, C., Charif, D., Latorre, A., Moya, A., Heddi, A., 2008. Identification of the weevil immune genes and their expression in the bacteriome tissue. BMC biology 6, 43.

Anselme, C., Vallier, A., Balmand, S., Fauvarque, M.O., Heddi, A., 2006. Host PGRP gene expression and bacterial release in endosymbiosis of the weevil *Sitophilus zeamais*. Applied and environmental microbiology 72, 6766-6772.

Arai, I., Ohta, M., Suzuki, A., Tanaka, S., Yoshizawa, Y., Sato, R., 2013. Immunohistochemical analysis of the role of hemocytin in nodule formation in the larvae of the silkworm, *Bombyx mori*. Journal of insect science 13, 125.

References

Aronstein, K.A., Murray, K.D., Saldivar, E., 2010. Transcriptional responses in honey bee larvae infected with chalkbrood fungus. Bmc Genomics 11, 391.

Ashok, Y., 2009. *Drosophila* toll pathway: the new model. Science signaling 2, jc1.

Asrat, S., de Jesus, D.A., Hempstead, A.D., Ramabhadran, V., Isberg, R.R., 2014. Bacterial pathogen manipulation of host membrane trafficking. Annual review of cell and developmental biology 30, 79-109.

Attardo, G.M., Lohs, C., Heddi, A., Alam, U.H., Yildirim, S., Aksoy, S., 2008. Analysis of milk gland structure and function in *Glossina morsitans*: milk protein production, symbiont populations and fecundity. Journal of insect physiology 54, 1236-1242.

Ballari, S., Farji-Brener, A.G., Tadey, M., 2007. Waste management in the leaf-cutting ant *Acromyrmex lobicornis*: division of labour, aggressive behaviour, and location of external refuse dumps. J. Insect Behav. 20, 87-98.

Balmand, S., Lohs, C., Aksoy, S., Heddi, A., 2013. Tissue distribution and transmission routes for the tsetse fly endosymbionts. Journal of invertebrate pathology 112 Suppl, S116-122.

Bang, K., Hwang, S., Lee, J., Cho, S., 2015. Identification of immunity-related genes in the larvae of *Protaetia brevitarsis seulensis* (Coleoptera: Cetoniidae) by a next-generation sequencing-based transcriptome analysis. Journal of insect science 15.

Barchuk, A.R., Bitondi, M.M., Simoes, Z.L., 2002. Effects of juvenile hormone and ecdysone on the timing of vitellogenin appearance in hemolymph of queen and worker pupae of *Apis mellifera*. Journal of insect science 2, 1.

Basbous, N., Coste, F., Leone, P., Vincentelli, R., Royet, J., Kellenberger, C., Roussel, A., 2011. The *Drosophila* peptidoglycan-recognition protein LF interacts with peptidoglycan-recognition protein LC to downregulate the Imd pathway. EMBO reports 12, 327-333.

Baumann, P., 2005. Biology of bacteriocyte-associated endosymbionts of plant sap-sucking insects. Annual review of microbiology 59, 155-189.

Bayles, K.W., Wesson, C.A., Liou, L.E., Fox, L.K., Bohach, G.A., Trumble, W.R., 1998. Intracellular *Staphylococcus aureus* escapes the endosome and induces apoptosis in epithelial cells. Infection and immunity 66, 336-342.

Berger, I., Fitzgerald, D.J., Richmond, T.J., 2004. Baculovirus expression system for heterologous multiprotein complexes. Nature biotechnology 22, 1583-1587.

Billen, J., 1985. Ultrastructure of the worker ovarioles in *Formica* ants. International Journal of Insect Morphololy and Embryology 14, 21-32.

Binari, R., Perrimon, N., 1994. Stripe-specific regulation of pair-rule genes by hopscotch, a putative Jak family tyrosine kinase in *Drosophila*. Genes & development 8, 300-312.

Bischoff, V., Vignal, C., Duvic, B., Boneca, I.G., Hoffmann, J.A., Royet, J., 2006. Downregulation of the *Drosophila* immune response by peptidoglycan-recognition proteins SC1 and SC2. PLoS pathogens 2, e14.

References

Blandin, S., Levashina, E.A., 2004. Thioester-containing proteins and insect immunity. Molecular immunology 40, 903-908.

Blandin, S., Shiao, S.H., Moita, L.F., Janse, C.J., Waters, A.P., Kafatos, F.C., Levashina, E.A., 2004. Complement-like protein TEP1 is a determinant of vectorial capacity in the malaria vector *Anopheles gambiae*. Cell 116, 661-670.

Blissard, G.W., 1996. Baculovirus-insect cell interactions. Cytotechnology 20, 73-93.

Blochmann, F., 1882. Über das Vorkommen bakterienähnlicher Gebilde in den Geweben und Eiern verschiedener Insekten. Zbl. Bakt 11, 234-240.

Bogomolovas, J., Simon, B., Sattler, M., Stier, G., 2009. Screening of fusion partners for high yield expression and purification of bioactive viscotoxins. Protein expression and purification 64, 16-23.

Bolton, B., 1995. A New and General Catalogue of the Ants of the World, Cambridge, MA, USA.

Bonasio, R., Zhang, G., Ye, C., Mutti, N.S., Fang, X., Qin, N., Donahue, G., Yang, P., Li, Q., Li, C., Zhang, P., Huang, Z., Berger, S.L., Reinberg, D., Wang, J., Liebig, J., 2010. Genomic comparison of the ants *Camponotus floridanus* and *Harpegnathos saltator*. Science 329, 1068-1071.

Bou Aoun, R., Hetru, C., Troxler, L., Doucet, D., Ferrandon, D., Matt, N., 2011. Analysis of thioester-containing proteins during the innate immune response of *Drosophila melanogaster*. Journal of innate immunity 3, 52-64.

Boulware, K.T., Daugherty, P.S., 2006. Protease specificity determination by using cellular libraries of peptide substrates (CLiPS). Proceedings of the National Academy of Sciences of the United States of America 103, 7583-7588.

Boutros, M., Agaisse, H., Perrimon, N., 2002. Sequential activation of signaling pathways during innate immune responses in *Drosophila*. Developmental cell 3, 711-722.

Braendle, C., Miura, T., Bickel, R., Shingleton, A.W., Kambhampati, S., Stern, D.L., 2003. Developmental Origin and Evolution of Bacteriocytes in the Aphid-*Buchnera* Symbiosis. Plos Biol 1, E21.

Brandt, A., 1874. Über die Eiröhren der *Blatta orientalis* (Periplaneta). Mem. Acad. Imp. Sci. 21, 1-30.

Brennan, C.A., Anderson, K.V., 2004. *Drosophila*: the genetics of innate immune recognition and response. Annual review of immunology 22, 457-483.

Brown, S., Hu, N., Hombria, J.C., 2001. Identification of the first invertebrate interleukin JAK/STAT receptor, the *Drosophila* gene domeless. Current biology : CB 11, 1700-1705.

Brun, S., Vidal, S., Spellman, P., Takahashi, K., Tricoire, H., Lemaitre, B., 2006. The MAPKKK Mekk1 regulates the expression of *Turandot* stress genes in response to septic injury in *Drosophila*. Genes to cells : devoted to molecular & cellular mechanisms 11, 397-407.

Buchner, P., 1918. Vergleichende Eistudien. I. die akzessorischen Kerne des Hymenoptereneies. Arch Mikroskop Anat II 90, 70-88.

References

Buchner, P., 1965. Endosymbiosis of Animals with Plant Microorganisms. Interscience, New York.

Bulet, P., Stocklin, R., 2005. Insect antimicrobial peptides: structures, properties and gene regulation. Protein and peptide letters 12, 3-11.

Bulet, P., Stocklin, R., Menin, L., 2004. Anti-microbial peptides: from invertebrates to vertebrates. Immunological reviews 198, 169-184.

Bulmer, M.S., Bachelet, I., Raman, R., Rosengaus, R.B., Sasisekharan, R., 2009. Targeting an antimicrobial effector function in insect immunity as a pest control strategy. Proceedings of the National Academy of Sciences of the United States of America 106, 12652-12657.

Büning, J., 2005. The telotrophic ovary known from *Neuropterida* exists also in the myxophagan beetle *Hydroscapha natans*. Development genes and evolution 215, 597-607.

Capurro, M.L., Moreira-Ferro, C.K., Marinotti, O., James, A.A., de Bianchi, A.G., 2000. Expression patterns of the larval and adult hexamerin genes of *Musca domestica*. Insect molecular biology 9, 169-177.

Cardoza, Y.J., Klebzig, K.D., Raffa, K.F., 2006. Bacteria in oral secretions of an endophytic insect inhibit antagonistic fungi. Ecol Entomol 31, 636-645.

Casteels-Josson, K., Capaci, T., Casteels, P., Tempst, P., 1993. Apidaecin multipeptide precursor structure: a putative mechanism for amplification of the insect antibacterial response. The EMBO journal 12, 1569-1578.

Casteels-Josson, K., Zhang, W., Capaci, T., Casteels, P., Tempst, P., 1994. Acute transcriptional response of the honeybee peptide-antibiotics gene repertoire and required post-translational conversion of the precursor structures. The Journal of biological chemistry 269, 28569-28575.

Casteels, P., Ampe, C., Jacobs, F., Tempst, P., 1993. Functional and chemical characterization of hymenoptaecin, an antibacterial polypeptide that is infection-inducible in the honeybee (*Apis mellifera*). The Journal of biological chemistry 268, 7044-7054.

Casteels, P., Ampe, C., Riviere, L., Van Damme, J., Elicone, C., Fleming, M., Jacobs, F., Tempst, P., 1990. Isolation and characterization of abaecin, a major antibacterial response peptide in the honeybee (*Apis mellifera*). European journal of biochemistry / FEBS 187, 381-386.

Cerenius, L., Söderhäll, K., 2004. The prophenoloxidase-activating system in invertebrates. Immunological reviews 198, 116-126.

Chang, C.I., Pili-Floury, S., Herve, M., Parquet, C., Chelliah, Y., Lemaitre, B., Mengin-Lecreulx, D., Deisenhofer, J., 2004. A *Drosophila* pattern recognition receptor contains a peptidoglycan docking groove and unusual L,D-carboxypeptidase activity. Plos Biol 2, E277.

Chapuisat, M., Oppliger, A., Magliano, P., Christe, P., 2007. Wood ants use resin to protect themselves against pathogens. Proceedings. Biological sciences / The Royal Society 274, 2013-2017.

Chen, H.W., Chen, X., Oh, S.W., Marinissen, M.J., Gutkind, J.S., Hou, S.X., 2002. *mom* identifies a receptor for the *Drosophila* JAK/STAT signal transduction pathway and encodes a protein distantly related to the mammalian cytokine receptor family. Genes & development 16, 388-398.

References

Chen, X., Li, S., Aksoy, S., 1999. Concordant evolution of a symbiont with its host insect species: molecular phylogeny of genus *Glossina* and its bacteriome-associated endosymbiont, *Wigglesworthia glossinidia*. Journal of molecular evolution 48, 49-58.

Choi, Y.S., Choo, Y.M., Lee, K.S., Yoon, H.J., Kim, I., Je, Y.H., Sohn, H.D., Jin, B.R., 2008. Cloning and expression profiling of four antibacterial peptide genes from the bumblebee *Bombus ignitus*. Comparative biochemistry and physiology. Part B, Biochemistry & molecular biology 150, 141-146.

Christe, P., Oppliger, A., Bancala, F., Castella, G., Chapuisat, M., 2003. Evidence for collective medication in ants. Ecol. Lett. 6, 19–22.

Clemens, D.L., Lee, B.Y., Horwitz, M.A., 2009. *Francisella tularensis* phagosomal escape does not require acidification of the phagosome. Infection and immunity 77, 1757-1773.

Colgan, T.J., Carolan, J.C., Bridgett, S.J., Sumner, S., Blaxter, M.L., Brown, M.J., 2011. Polyphenism in social insects: insights from a transcriptome-wide analysis of gene expression in the life stages of the key pollinator, *Bombus terrestris*. Bmc Genomics 12, 623.

Cornet, B., Bonmatin, J.M., Hetru, C., Hoffmann, J.A., Ptak, M., Vovelle, F., 1995. Refined three-dimensional solution structure of insect defensin A. Structure 3, 435-448.

Correia, J.J., Detrich, H.W., 2009. Biophysical Tools for Biologists: *In Vivo* Techniques. Academic Press.

Cox, J., Hein, M.Y., Luber, C.A., Paron, I., Nagaraj, N., Mann, M., 2014. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. Molecular & cellular proteomics : MCP 13, 2513-2526.

Cox, J., Mann, M., 2008. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nature biotechnology 26, 1367-1372.

Cremer, S., Armitage, S.A., Schmid-Hempel, P., 2007. Social immunity. Current biology : CB 17, R693-702.

Currie, C.R., Wong, B., Stuart, A.E., Schultz, T.R., Rehner, S.A., Mueller, U.G., Sung, G.H., Spatafora, J.W., Straus, N.A., 2003. Ancient tripartite coevolution in the attine ant-microbe symbiosis. Science 299, 386-388.

Dale, C., Moran, N.A., 2006. Molecular interactions between bacterial symbionts and their hosts. Cell 126, 453-465.

Davies, S.A., Overend, G., Sebastian, S., Cundall, M., Cabrero, P., Dow, J.A., Terhzaz, S., 2012. Immune and stress response 'cross-talk' in the *Drosophila* Malpighian tubule. Journal of insect physiology 58, 488-497.

De Bary, A., 1879. Die Erscheinung der Symbiose. Straßburg: Verlag v. K. J. Trübner.

De Gregorio, E., Han, S.J., Lee, W.J., Baek, M.J., Osaki, T., Kawabata, S., Lee, B.L., Iwanaga, S., Lemaitre, B., Brey, P.T., 2002. An immune-responsive Serpin regulates the melanization cascade in *Drosophila*. Developmental cell 3, 581-592.

References

De Gregorio, E., Spellman, P.T., Rubin, G.M., Lemaitre, B., 2001. Genome-wide analysis of the *Drosophila* immune response by using oligonucleotide microarrays. Proceedings of the National Academy of Sciences of the United States of America 98, 12590-12595.

de Morais Guedes, S., Vitorino, R., Domingues, R., Tomer, K., Correia, A.J., Amado, F., Domingues, P., 2005. Proteomics of immune-challenged *Drosophila melanogaster* larvae hemolymph. Biochemical and biophysical research communications 328, 106-115.

de Souza, D.J., Bezier, A., Depoix, D., Drezen, J.M., Lenoir, A., 2009. *Blochmannia* endosymbionts improve colony growth and immune defence in the ant *Camponotus fellah*. BMC microbiology 9, 29.

de Souza, D.J., Van Vlaenderen, J., Moret, Y., Lenoir, A., 2008. Immune response affects ant trophallactic behaviour. Journal of insect physiology 54, 828-832.

Degnan, P.H., Lazarus, A.B., Brock, C.D., Wernegreen, J.J., 2004. Host–Symbiont Stability and Fast Evolutionary Rates in an Ant–Bacterium Association: Cospeciation of *Camponotus* Species and Their Endosymbionts, *Candidatus* Blochmannia. Systematic biology 53, 95-110.

Dmochowska, A., Dignard, D., Henning, D., Thomas, D.Y., Bussey, H., 1987. Yeast *KEX1* gene encodes a putative protease with a carboxypeptidase B-like function involved in killer toxin and alpha-factor precursor processing. Cell 50, 573-584.

do Amaral, J.B., Machado-Santelli, G.M., 2009. Three-dimensional Reconstruction of Ovaries of Leaf-cutting Ant (*Atta sexdens rubropilosa*) Queens (Hymenoptera: Formicidae). SOCIOBIOLOGY 53, 379-388.

Doherty, G.J., McMahon, H.T., 2009. Mechanisms of Endocytosis. Annual review of biochemistry 78, 857-902.

Dostert, C., Jouanguy, E., Irving, P., Troxler, L., Galiana-Arnoux, D., Hetru, C., Hoffmann, J.A., Imler, J.L., 2005. The Jak-STAT signaling pathway is required but not sufficient for the antiviral response of *Drosophila*. Nature immunology 6, 946-953.

Douglas, A.E., 1998. Nutritional Interactions in Insect-Microbial Symbioses: Aphids and Their Symbiotic Bacteria *Buchnera*. Annual Review of Entomology 43, 17–37.

Du, Z.Q., Ren, Q., Zhao, X.F., Wang, J.X., 2009. A double WAP domain (DWD)-containing protein with proteinase inhibitory activity in Chinese white shrimp, Fenneropenaeus chinensis. Comparative biochemistry and physiology. Part B, Biochemistry & molecular biology 154, 203-210.

Dunbar, H.E., Wilson, A.C.C., Ferguson, N.R., Moran, N.A., 2007. Aphid thermal tolerance is governed by a point mutation in bacterial symbionts. Plos Biol 5, 1006-1015.

Ehrhardt, C., Schmolke, M., Matzke, A., Knoblauch, A., Will, C., Wixler, V., Ludwig, S., 2006. Polyethylenimine, a cost-effective transfection reagent. Signal Transduction 6, 179-184.

El Chamy, L., Leclerc, V., Caldelari, I., Reichhart, J.M., 2008. Sensing of 'danger signals' and pathogen-associated molecular patterns defines binary signaling pathways 'upstream' of Toll. Nature immunology 9, 1165-1170.

References

Endler, A., Liebig, J., Schmitt, T., Parker, J.E., Jones, G.R., Schreier, P., Holldobler, B., 2004. Surface hydrocarbons of queen eggs regulate worker reproduction in a social insect. Proceedings of the National Academy of Sciences of the United States of America 101, 2945-2950.

Epsky, N.D., Capinera, J.L., 1988. Efficacy of the entomogenous nematode *Steinernema feltiae* against a subterranean termite, *Reticulitermes tibialis* (Isoptera, Rhinotermitidae). . J. Econ. Entomol. 81, 1313–1317.

Evans, J.D., Aronstein, K., Chen, Y.P., Hetru, C., Imler, J.L., Jiang, H., Kanost, M., Thompson, G.J., Zou, Z., Hultmark, D., 2006. Immune pathways and defence mechanisms in honey bees *Apis mellifera*. Insect molecular biology 15, 645-656.

Fang, S.-M., 2012. Insect glutathione S-transferase: a review of comparative genomic studies and response to xenobiotics. Bulletin of Insectology 65, 265-271.

Feldhaar, H., 2011. Bacterial symbionts as mediators of ecologically important traits of insect hosts. Ecol Entomol 36, 533-543.

Feldhaar, H., Gross, R., 2008. Immune reactions of insects on bacterial pathogens and mutualists. Microbes and infection / Institut Pasteur 10, 1082-1088.

Feldhaar, H., Gross, R., 2009. Insects as hosts for mutualistic bacteria. International journal of medical microbiology : IJMM 299, 1-8.

Feldhaar, H., Straka, J., Krischke, M., Berthold, K., Stoll, S., Mueller, M.J., Gross, R., 2007. Nutritional upgrading for omnivorous carpenter ants by the endosymbiont *Blochmannia*. BMC biology 5, 48.

Fellous, S., Lazzaro, B.P., 2011. Potential for evolutionary coupling and decoupling of larval and adult immune gene expression. Molecular ecology 20, 1558-1567.

Feng, M., Ramadan, H., Han, B., Fang, Y., Li, J., 2014. Hemolymph proteome changes during worker brood development match the biological divergences between western honey bees (*Apis mellifera*) and eastern honey bees (*Apis cerana*). Bmc Genomics 15, 563.

Fernandez-Marin, H., Zimmerman, J.K., Rehner, S.A., Wcislo, W.T., 2006. Active use of the metapleural glands by ants in controlling fungal infection. Proceedings. Biological sciences / The Royal Society 273, 1689-1695.

Fernandez, N.Q., Grosshans, J., Goltz, J.S., Stein, D., 2001. Separable and redundant regulatory determinants in Cactus mediate its dorsal group dependent degradation. Development 128, 2963-2974.

Finn, R.D., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J.E., Gavin, O.L., Gunasekaran, P., Ceric, G., Forslund, K., Holm, L., Sonnhammer, E.L., Eddy, S.R., Bateman, A., 2010. The Pfam protein families database. Nucleic acids research 38, D211-222.

Finn, R.N., 2007. Vertebrate yolk complexes and the functional implications of phosvitins and other subdomains in vitellogenins. Biology of reproduction 76, 926-935.

Fitzgerald, D.J., Berger, P., Schaffitzel, C., Yamada, K., Richmond, T.J., Berger, I., 2006. Protein complex expression by using multigene baculoviral vectors. Nature methods 3, 1021-1032.

References

Fjell, C.D., Hancock, R.E., Cherkasov, A., 2007. AMPer: a database and an automated discovery tool for antimicrobial peptides. Bioinformatics 23, 1148-1155.

Fleig, R., 1995. Role of the follicle cells for yolk uptake in ovarian follicles of the honey bee *Apis mellifera* L. (Hymenoptera : Apidae). lnt. J. Insect Morphol. & Embryol. 24, 427-433.

Fricker, L.D., 2012. Neuropeptides and other bioactive peptides:from discovery to function. Colloq Ser Neuropept 3, 1–122.

Fujiwara, K., Taguchi, H., 2007. Filamentous morphology in GroE-depleted *Escherichia coli* induced by impaired folding of FtsE. Journal of bacteriology 189, 5860-5866.

Fuller, R.S., Brake, A., Thorner, J., 1989. Yeast prohormone processing enzyme (KEX2 gene product) is a Ca2+-dependent serine protease. Proceedings of the National Academy of Sciences of the United States of America 86, 1434-1438.

Futschik, M.E., Carlisle, B., 2005. Noise-robust soft clustering of gene expression time-course data. Journal of bioinformatics and computational biology 3, 965-988.

Gandhe, A.S., Arunkumar, K.P., John, S.H., Nagaraju, J., 2006. Analysis of bacteria-challenged wild silkmoth, *Antheraea mylitta* (lepidoptera) transcriptome reveals potential immune genes. Bmc Genomics 7, 184.

Ganesan, S., Aggarwal, K., Paquette, N., Silverman, N., 2011. NF-kappaB/Rel proteins and the humoral immune responses of *Drosophila melanogaster*. Current topics in microbiology and immunology 349, 25-60.

Gao, B., Zhu, S., 2010. Characterization of a hymenoptaecin-like antimicrobial peptide in the parasitic wasp *Nasonia vitripennis*. Process Biochemistry 45, 139-146.

Gasparian, M.E., Bychkov, M.L., Dolgikh, D.A., Kirpichnikov, M.P., 2011. Strategy for improvement of enteropeptidase efficiency in tag removal processes. Protein expression and purification 79, 191-196.

Geier, M., 2009. Experimentelle Ansiedlung von Pathogenen im Darm von *Camponotus floridanus*. Diploma Thesis, Universität Würzburg.

Geiser, D.L., Chavez, C.A., Flores-Munguia, R., Winzerling, J.J., Pham, D.Q., 2003. *Aedes aegypti* ferritin. European journal of biochemistry / FEBS 270, 3667-3674.

Gerardo, N.M., Altincicek, B., Anselme, C., Atamian, H., Barribeau, S.M., de Vos, M., Duncan, E.J., Evans, J.D., Gabaldon, T., Ghanim, M., Heddi, A., Kaloshian, I., Latorre, A., Moya, A., Nakabachi, A., Parker, B.J., Perez-Brocal, V., Pignatelli, M., Rahbe, Y., Ramsey, J.S., Spragg, C.J., Tamames, J., Tamarit, D., Tamborindeguy, C., Vincent-Monegat, C., Vilcinskas, A., 2010. Immunity and other defenses in pea aphids, *Acyrthosiphon pisum*. Genome biology 11, R21.

Gil, R., Latorre, A., Moya, A., 2004. Bacterial endosymbionts of insects: insights from comparative genomics. Environmental microbiology 6, 1109-1122.

Gil, R., Silva, F.J., Zientz, E., Delmotte, F., Gonzalez-Candelas, F., Latorre, A., Rausell, C., Kamerbeek, J., Gadau, J., Holldobler, B., van Ham, R.C., Gross, R., Moya, A., 2003. The genome

sequence of *Blochmannia floridanus*: Comparative analysis of reduced genomes. Proceedings of the National Academy of Sciences of the United States of America 100, 9388-9393.

Gilliam, M., Taber, S., III, Lorenz, B.J., Prest, D.B., 1988. Factors affecting development of chalkbrood disease in colonies of honey bees, *Apis mellifera,* fed pollen contaminated with *Ascosphaera apis*. J. Inv. Pathol. 52, 314–325.

Gobert, V., Gottar, M., Matskevich, A.A., Rutschmann, S., Royet, J., Belvin, M., Hoffmann, J.A., Ferrandon, D., 2003. Dual activation of the *Drosophila* toll pathway by two pattern recognition receptors. Science 302, 2126-2130.

Gonzalez-Reyes, A., Elliott, H., St Johnston, D., 1997. Oocyte determination and the origin of polarity in *Drosophila*: the role of the *spindle* genes. Development 124, 4927-4937.

Gorman, M.J., An, C., Kanost, M.R., 2007. Characterization of tyrosine hydroxylase from *Manduca sexta*. Insect biochemistry and molecular biology 37, 1327-1337.

Gottar, M., Gobert, V., Matskevich, A.A., Reichhart, J.M., Wang, C., Butt, T.M., Belvin, M., Hoffmann, J.A., Ferrandon, D., 2006. Dual detection of fungal infections in *Drosophila* via recognition of glucans and sensing of virulence factors. Cell 127, 1425-1437.

Gotz, P., Weise, C., Kopacek, P., Losen, S., Wiesner, A., 1997. Isolated Apolipophorin III from *Galleria mellonella* stimulates the Immune Reactions of this insect. Journal of insect physiology 43, 383-391.

Grell, R.F., Generoso, E.E., 1982. A temporal study at the ultrastructural level of the developing pro-oocyte of *Drosophila melanogaster*. Chromosoma 87, 49-75.

Gross, J., 1903. Untersuchungen über die Histologie des Insecten Ovariums. Zool. Jb. Abt. Morphol. 18, 71-186.

Gross, R., Vavre, F., Heddi, A., Hurst, G.D., Zchori-Fein, E., Bourtzis, K., 2009. Immunity and symbiosis. Molecular microbiology 73, 751-759.

Gupta, L., Noh, J.Y., Jo, Y.H., Oh, S.H., Kumar, S., Noh, M.Y., Lee, Y.S., Cha, S.J., Seo, S.J., Kim, I., Han, Y.S., Barillas-Mury, C., 2010. Apolipophorin-III mediates antiplasmodial epithelial responses in *Anopheles gambiae* (G3) mosquitoes. PloS one 5, e15410.

Gupta, S.K., 2016. Re-annotation of *Camponotus floridanus* genome and characterization of innate immunity transcriptome response responses to bacterial infections.

Gupta, S.K., Kupper, M., Ratzka, C., Feldhaar, H., Vilcinskas, A., Gross, R., Dandekar, T., Forster, F., 2015. Scrutinizing the immune defence inventory of *Camponotus floridanus* applying total transcriptome sequencing. Bmc Genomics 16.

Haas, A., 2007. The phagosome: compartment with a license to kill. Traffic 8, 311-330.

Haine, E.R., Moret, Y., Siva-Jothy, M.T., Rolff, J., 2008. Antimicrobial defense and persistent infection in insects. Science 322, 1257-1259.

References

Halwani, A., Dunphy, G., 1997. Haemolymph proteins of larvae of *Galleria mellonella* detoxify endotoxins of the insect pathogenic bacteria *Xenorhabdus nematophilus* (Enterobacteriaceae). Journal of insect physiology 43, 1023-1029.

Halwani, A.E., Niven, D.F., Dunphy, G.B., 2000. Apolipophorin-III and the interactions of lipoteichoic acids with the immediate immune responses of *Galleria mellonella*. Journal of invertebrate pathology 76, 233-241.

Hamilton, C., Lejeune, B.T., Rosengaus, R.B., 2011. Trophallaxis and prophylaxis: social immunity in the carpenter ant *Camponotus pennsylvanicus*. Biology letters 7, 89-92.

Hamilton, P.T., Perlman, S.J., 2013. Host defense via symbiosis in *Drosophila*. PLoS pathogens 9, e1003808.

Harris, H.L., Brennan, L.J., Keddie, B.A., Braig, H.R., 2010. Bacterial symbionts in insects: balancing life and death. Symbiosis 51, 37-53.

Harrison, D.A., McCoon, P.E., Binari, R., Gilman, M., Perrimon, N., 1998. *Drosophila* unpaired encodes a secreted protein that activates the JAK signaling pathway. Genes & development 12, 3252-3263.

Hart, A.G., Ratnieks, F.L.W., 2001. Task partitioning, division of labour and nest compartimentalisation collectively isolate hazardous waste in the leafcutting ant *Atta cephalotes.* Behav. Ecol. Sociobiol. 49, 387–392.

Hashimoto, Y., Tabuchi, Y., Sakurai, K., Kutsuna, M., Kurokawa, K., Awasaki, T., Sekimizu, K., Nakanishi, Y., Shiratsuchi, A., 2009. Identification of lipoteichoic acid as a ligand for draper in the phagocytosis of *Staphylococcus aureus* by *Drosophila* hemocytes. Journal of immunology 183, 7451-7460.

Hecht, O., 1923. Embryonalentwicklung und Symbiose bei *Camponotus ligniperda*. Zeitschrift für Wissenschaftliche Zoologie 122, 173–204.

Heddi, A., Vallier, A., Anselme, C., Xin, H., Rahbe, Y., Wackers, F., 2005. Molecular and cellular profiles of insect bacteriocytes: mutualism and harm at the initial evolutionary step of symbiogenesis. Cellular microbiology 7, 293-305.

Hedges, L.M., Brownlie, J.C., O'Neill, S.L., Johnson, K.N., 2008. *Wolbachia* and virus protection in insects. Science 322, 702.

Hericourt, F., Blanc, S., Redeker, V., Jupin, I., 2000. Evidence for phosphorylation and ubiquitinylation of the turnip yellow mosaic virus RNA-dependent RNA polymerase domain expressed in a baculovirus-insect cell system. The Biochemical journal 349, 417-425.

Herren, J.K., Lemaitre, B., 2011. *Spiroplasma* and host immunity: activation of humoral immune responses increases endosymbiont load and susceptibility to certain Gram-negative bacterial pathogens in *Drosophila melanogaster*. Cellular microbiology 13, 1385-1396.

Hetru, C., Hoffmann, J.A., 2009. NF-kappaB in the immune response of *Drosophila*. Cold Spring Harbor perspectives in biology 1, a000232.

References

Hodder, A.N., Crewther, P.E., Matthew, M.L., Reid, G.E., Moritz, R.L., Simpson, R.J., Anders, R.F., 1996. The disulfide bond structure of Plasmodium apical membrane antigen-1. The Journal of biological chemistry 271, 29446-29452.

Hoffmann, J.A., 2003. The immune response of *Drosophila*. Nature 426, 33-38.

Hölldobler, B., Engel-Siegel, H., 1984. On the metapleural gland of ants. Psyche 91, 201-224.

Hombria, J.C., Brown, S., Hader, S., Zeidler, M.P., 2005. Characterisation of Upd2, a *Drosophila* JAK/STAT pathway ligand. Developmental biology 288, 420-433.

Hsu, T.A., Betenbaugh, M.J., 1997. Coexpression of molecular chaperone BiP improves immunoglobulin solubility and IgG secretion from Trichoplusia ni insect cells. Biotechnology progress 13, 96-104.

Huang, L., Ching, C.B., Jiang, R., Leong, S.S., 2008. Production of bioactive human beta-defensin 5 and 6 in *Escherichia coli* by soluble fusion expression. Protein expression and purification 61, 168-174.

Hurst, G.D., Anbutsu, H., Kutsukake, M., Fukatsu, T., 2003. Hidden from the host: *Spiroplasma* bacteria infecting *Drosophila* do not cause an immune response, but are suppressed by ectopic immune activation. Insect molecular biology 12, 93-97.

Hussain, M., Asgari, S., 2014. MicroRNAs as mediators of insect host-pathogen interactions and immunity. Journal of insect physiology 70, 151-158.

Hussain, M., Frentiu, F.D., Moreira, L.A., O'Neill, S.L., Asgari, S., 2011. *Wolbachia* uses host microRNAs to manipulate host gene expression and facilitate colonization of the dengue vector *Aedes aegypti*. Proceedings of the National Academy of Sciences of the United States of America 108, 9250-9255.

Hutchence, K.J., Fischer, B., Paterson, S., Hurst, G.D., 2011. How do insects react to novel inherited symbionts? A microarray analysis of *Drosophila melanogaster* response to the presence of natural and introduced *Spiroplasma*. Molecular ecology 20, 950-958.

Hybiske, K., Stephens, R.S., 2008. Exit strategies of intracellular pathogens. Nature reviews. Microbiology 6, 99-110.

Ip, Y.T., Davis, R.J., 1998. Signal transduction by the c-Jun N-terminal kinase (JNK)-from inflammation to development. Current opinion in cell biology 10, 205-219.

James, D.C., Freedman, R.B., Hoare, M., Ogonah, O.W., Rooney, B.C., Larionov, O.A., Dobrovolsky, V.N., Lagutin, O.V., Jenkins, N., 1995. N-glycosylation of recombinant human interferon-gamma produced in different animal expression systems. Bio/technology 13, 592-596.

Jomori, T., Natori, S., 1992. Function of the lipopolysaccharide-binding protein of *Periplaneta americana* as an opsonin. FEBS letters 296, 283-286.

Kaltenpoth, M., 2009. Actinobacteria as mutualists: general healthcare for insects? Trends in microbiology 17, 529-535.

References

Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., Katayama, T., Kawashima, S., Okuda, S., Tokimatsu, T., Yamanishi, Y., 2008. KEGG for linking genomes to life and the environment. Nucleic acids research 36, D480-484.

Kaneko, T., Yano, T., Aggarwal, K., Lim, J.H., Ueda, K., Oshima, Y., Peach, C., Erturk-Hasdemir, D., Goldman, W.E., Oh, B.H., Kurata, S., Silverman, N., 2006. PGRP-LC and PGRP-LE have essential yet distinct functions in the *Drosophila* immune response to monomeric DAP-type peptidoglycan. Nature immunology 7, 715-723.

Kanost, M.R., Jiang, H., Yu, X.Q., 2004. Innate immune responses of a lepidopteran insect, *Manduca sexta*. Immunological reviews 198, 97-105.

Kapust, R.B., Tozser, J., Copeland, T.D., Waugh, D.S., 2002. The P1' specificity of tobacco etch virus protease. Biochemical and biophysical research communications 294, 949-955.

Karlsson, C., Korayem, A.M., Scherfer, C., Loseva, O., Dushay, M.S., Theopold, U., 2004. Proteomic analysis of the *Drosophila* larval hemolymph clot. The Journal of biological chemistry 279, 52033-52041.

Khila, A., Abouheif, E., 2008. Reproductive constraint is a developmental mechanism that maintains social harmony in advanced ant societies. Proceedings of the National Academy of Sciences of the United States of America 105, 17884-17889.

Kikuchi, Y., 2009. Endosymbiotic bacteria in insects: their diversity and culturability. Microbes and environments / JSME 24, 195-204.

King, R.C., Büning, J., 1985. The origin and functioning of insect oocytes and nurse cells, In: Kerkut, G.A., Gilbert, L.J. (Eds.), Comprehensive insect physiology, biochemistry and pharmacology: embryogenesis and reproduction. Pergamon Press, Oxford, pp. 37-82.

Kingsolver, M.B., Huang, Z., Hardy, R.W., 2013. Insect antiviral innate immunity: pathways, effectors, and connections. Journal of molecular biology 425, 4921-4936.

Klag, J., Bilinski, S., 1994. Germ cell cluster formation and oogenesis in the hymenopteran *Coleocentrotus soldanskii*. Tissue & cell 26, 699-706.

Klasson, L., Andersson, S.G., 2010. Research on small genomes: Implications for synthetic biology. BioEssays : news and reviews in molecular, cellular and developmental biology 32, 288-295.

Klein, A., Schrader, L., Gil, R., Manzano-Marin, A., Florez, L., Wheeler, D., Werren, J.H., Latorre, A., Heinze, J., Kaltenpoth, M., Moya, A., Oettler, J., 2016. A novel intracellular mutualistic bacterium in the invasive ant *Cardiocondyla obscurior*. The ISME journal 10, 376-388.

Kleineidam, C.J., Obermayer, M., Halbich, W., Rossler, W., 2005. A macroglomerulus in the antennal lobe of leaf-cutting ant workers and its possible functional significance. Chemical senses 30, 383-392.

Kleino, A., Silverman, N., 2014. The *Drosophila* IMD pathway in the activation of the humoral immune response. Developmental and comparative immunology 42, 25-35.

References

Kobialka, M., Michalik, A., Walczak, M., Junkiert, L., Szklarzewicz, T., 2015. *Sulcia* symbiont of the leafhopper *Macrosteles laevis* (Ribaut, 1927) (Insecta, Hemiptera, Cicadellidae: Deltocephalinae) harbors *Arsenophonus* bacteria. Protoplasma.

Koch, E.A., Smith, P.A., King, R.C., 1967. The division and differentiation of *Drosophila* cystocytes. Journal of morphology 121, 55-70.

Kocks, C., Cho, J.H., Nehme, N., Ulvila, J., Pearson, A.M., Meister, M., Strom, C., Conto, S.L., Hetru, C., Stuart, L.M., Stehle, T., Hoffmann, J.A., Reichhart, J.M., Ferrandon, D., Ramet, M., Ezekowitz, R.A., 2005. Eater, a transmembrane protein mediating phagocytosis of bacterial pathogens in *Drosophila*. Cell 123, 335-346.

Koga, R., Meng, X.Y., Tsuchida, T., Fukatsu, T., 2012. Cellular mechanism for selective vertical transmission of an obligate insect symbiont at the bacteriocyte-embryo interface. Proceedings of the National Academy of Sciences of the United States of America 109, E1230-1237.

Koizumi, N., Imamura, M., Kadotani, T., Yaoi, K., Iwahana, H., Sato, R., 1999. The lipopolysaccharide-binding protein participating in hemocyte nodule formation in the silkworm *Bombyx mori* is a novel member of the C-type lectin superfamily with two different tandem carbohydrate-recognition domains. FEBS letters 443, 139-143.

Kolliopoulou, A., Van Nieuwerburgh, F., Stravopodis, D.J., Deforce, D., Swevers, L., Smagghe, G., 2015. Transcriptome Analysis of *Bombyx mori* Larval Midgut during Persistent and Pathogenic Cytoplasmic Polyhedrosis Virus Infection. PloS one 10, e0121447.

Konrad, M., Vyleta, M.L., Theis, F.J., Stock, M., Tragust, S., Klatt, M., Drescher, V., Marr, C., Ugelvig, L.V., Cremer, S., 2012. Social transfer of pathogenic fungus promotes active immunisation in ant colonies. Plos Biol 10, e1001300.

Kostaropoulos, I., Papadopoulos, A.I., Metaxakis, A., Boukouvala, E., Papadopoulou-Mourkidou, E., 2001a. Glutathione S-transferase in the defence against pyrethroids in insects. Insect biochemistry and molecular biology 31, 313-319.

Kostaropoulos, I., Papadopoulos, A.I., Metaxakis, A., Boukouvala, E., Papadopoulou-Mourkidou, E., 2001b. The role of glutathione S-transferases in the detoxification of some organophosphorus insecticides in larvae and pupae of the yellow mealworm, *Tenebrio molitor* (Coleoptera: Tenebrionidae). Pest management science 57, 501-508.

Kotani, E., Yamakawa, M., Iwamoto, S., Tashiro, M., Mori, H., Sumida, M., Matsubara, F., Taniai, K., Kadono-Okuda, K., Kato, Y., Mori, H., 1994. Cloning and expression of the gene of hemocytin, an insect humoral lectin which is homologous with the mammalian von Willebrand factor. Biochimica et biophysica acta, 245-258.

Kounatidis, I., Ligoxygakis, P., 2012. *Drosophila* as a model system to unravel the layers of innate immunity to infection. Open biology 2, 120075.

Kramm, K.R., West, D.F., Rockenbach, P.G., 1982. Termite pathogens: transfer of the entomopathogen *Metarhizium anisopliae* between *Reticulitermes* sp. termites. J. Inv. Pathol. 40, 1–6.

Kremer, J.R., Mastronarde, D.N., McIntosh, J.R., 1996. Computer visualization of three-dimensional image data using IMOD. Journal of structural biology 116, 71-76.

References

Kupper, M., 2011. Heterologe Expression und Characterisierung eines *Camponotus floridanus* Hymenoptaecins und Herstellung eines spezifischen *Blochmannia floridanus* Antiserums. Universität Würzburg.

Kurata, S., 2010. Extracellular and intracellular pathogen recognition by *Drosophila* PGRP-LE and PGRP-LC. International immunology 22, 143-148.

Kurata, S., 2014. Peptidoglycan recognition proteins in *Drosophila* immunity. Developmental and comparative immunology 42, 36-41.

Kurland, C., Gallant, J., 1996. Errors of heterologous protein expression. Current Opinion Biotechnology 7, 489-493.

Kurucz, E., Markus, R., Zsamboki, J., Folkl-Medzihradszky, K., Darula, Z., Vilmos, P., Udvardy, A., Krausz, I., Lukacsovich, T., Gateff, E., Zettervall, C.J., Hultmark, D., Ando, I., 2007. Nimrod, a putative phagocytosis receptor with EGF repeats in *Drosophila* plasmatocytes. Current biology : CB 17, 649-654.

Lanot, R., Zachary, D., Holder, F., Meister, M., 2001. Postembryonic hematopoiesis in *Drosophila*. Developmental biology 230, 243-257.

LaVallie, E.R., DiBlasio, E.A., Kovacic, S., Grant, K.L., Schendel, P.F., McCoy, J.M., 1993. A thioredoxin gene fusion expression system that circumvents inclusion body formation in the *E. coli* cytoplasm. Bio/technology 11, 187-193.

Lavine, M.D., Strand, M.R., 2002. Insect hemocytes and their role in immunity. Insect biochemistry and molecular biology 32, 1295-1309.

Le Conte, Y., Alaux, C., Martin, J.F., Harbo, J.R., Harris, J.W., Dantec, C., Severac, D., Cros-Arteil, S., Navajas, M., 2011. Social immunity in honeybees (*Apis mellifera*): transcriptome analysis of varroa-hygienic behaviour. Insect molecular biology 20, 399-408.

Lee, M.H., Osaki, T., Lee, J.Y., Baek, M.J., Zhang, R., Park, J.W., Kawabata, S., Soderhall, K., Lee, B.L., 2004. Peptidoglycan recognition proteins involved in 1,3-beta-D-glucan-dependent prophenoloxidase activation system of insect. The Journal of biological chemistry 279, 3218-3227.

Lemaitre, B., Hoffmann, J., 2007. The host defense of *Drosophila melanogaster*. Annual review of immunology 25, 697-743.

Leon, L.J., Idangodage, H., Wan, C.P., Weers, P.M., 2006. Apolipophorin III: lipopolysaccharide binding requires helix bundle opening. Biochemical and biophysical research communications 348, 1328-1333.

Levashina, E.A., Moita, L.F., Blandin, S., Vriend, G., Lagueux, M., Kafatos, F.C., 2001. Conserved role of a complement-like protein in phagocytosis revealed by dsRNA knockout in cultured cells of the mosquito, *Anopheles gambiae*. Cell 104, 709-718.

Levy, F., Bulet, P., Ehret-Sabatier, L., 2004. Proteomic analysis of the systemic immune response of Drosophila. Molecular & cellular proteomics : MCP 3, 156-166.

References

Lhocine, N., Ribeiro, P.S., Buchon, N., Wepf, A., Wilson, R., Tenev, T., Lemaitre, B., Gstaiger, M., Meier, P., Leulier, F., 2008. PIMS Modulates Immune Tolerance by Negatively Regulating *Drosophila* Innate Immune Signaling. Cell host & microbe 4, 147-158.

Li, D., Liang, Y., Wang, X., Wang, L., Qi, M., Yu, Y., Luan, Y., 2015. Transcriptomic Analysis of *Musca domestica* to Reveal Key Genes of the Prophenoloxidase-Activating System. G3 5, 1827-1841.

Li, F., Wang, L., Qiu, L., Zhang, H., Gai, Y., Song, L., 2012. A double WAP domain-containing protein from Chinese mitten crab Eriocheir sinensis with antimicrobial activities against Gram-negative bacteria and yeast. Developmental and comparative immunology 36, 183-190.

Li, Y., 2011. Recombinant production of antimicrobial peptides in *Escherichia coli*: a review. Protein expression and purification 80, 260-267.

Li, Z., Zhang, S., Liu, Q., 2008. Vitellogenin functions as a multivalent pattern recognition receptor with an opsonic activity. PloS one 3, e1940.

Li, Z., Zhang, S., Zhang, J., Liu, M., Liu, Z., 2009. Vitellogenin is a cidal factor capable of killing bacteria via interaction with lipopolysaccharide and lipoteichoic acid. Molecular immunology 46, 3232-3239.

Ligoxygakis, P., Pelte, N., Hoffmann, J.A., Reichhart, J.M., 2002. Activation of *Drosophila* Toll during fungal infection by a blood serine protease. Science 297, 114-116.

Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25, 402-408.

Login, F.H., Balmand, S., Vallier, A., Vincent-Monegat, C., Vigneron, A., Weiss-Gayet, M., Rochat, D., Heddi, A., 2011. Antimicrobial peptides keep insect endosymbionts under control. Science 334, 362-365.

Login, F.H., Heddi, A., 2013. Insect immune system maintains long-term resident bacteria through a local response. Journal of insect physiology 59, 232-239.

Lourenco, A.P., Guidugli-Lazzarini, K.R., Freitas, F.C., Bitondi, M.M., Simoes, Z.L., 2013. Bacterial infection activates the immune system response and dysregulates microRNA expression in honey bees. Insect biochemistry and molecular biology 43, 474-482.

Lourenco, A.P., Martins, J.R., Bitondi, M.M., Simoes, Z.L., 2009. Trade-off between immune stimulation and expression of storage protein genes. Archives of insect biochemistry and physiology 71, 70-87.

Lourenco, A.P., Martins, J.R., Guidugli-Lazzarini, K.R., Macedo, L.M., Bitondi, M.M., Simoes, Z.L., 2012. Potential costs of bacterial infection on storage protein gene expression and reproduction in queenless *Apis mellifera* worker bees on distinct dietary regimes. Journal of insect physiology 58, 1217-1225.

Ma, G., Hay, D., Li, D., Asgari, S., Schmidt, O., 2006. Recognition and inactivation of LPS by lipophorin particles. Developmental and comparative immunology 30, 619-626.

References

Martinez, T., Burmester, T., Veenstra, J.A., Wheeler, D., 2000. Sequence and evolution of a hexamerin from the ant *Camponotus festinatus*. Insect molecular biology 9, 427-431.

Martins, J.R., Nunes, F.M., Cristino, A.S., Simoes, Z.L., Bitondi, M.M., 2010. The four hexamerin genes in the honey bee: structure, molecular evolution and function deduced from expression patterns in queens, workers and drones. BMC molecular biology 11, 23.

Mastronarde, D.N., 2005. Automated electron microscope tomography using robust prediction of specimen movements. Journal of structural biology 152, 36-51.

Matsuura, Y., Kikuchi, Y., Hosokawa, T., Koga, R., Meng, X.Y., Kamagata, Y., Nikoh, N., Fukatsu, T., 2012a. Evolution of symbiotic organs and endosymbionts in lygaeid stinkbugs. The ISME journal 6, 397-409.

Matsuura, Y., Kikuchi, Y., Meng, X.Y., Koga, R., Fukatsu, T., 2012b. Novel clade of alphaproteobacterial endosymbionts associated with stinkbugs and other arthropods. Applied and environmental microbiology 78, 4149-4156.

McCutcheon, J.P., Moran, N.A., 2012. Extreme genome reduction in symbiotic bacteria. Nature reviews. Microbiology 10, 13-26.

Meng, Q., Yu, H.Y., Zhang, H., Zhu, W., Wang, M.L., Zhang, J.H., Zhou, G.L., Li, X., Qin, Q.L., Hu, S.N., Zou, Z., 2015. Transcriptomic insight into the immune defenses in the ghost moth, *Hepialus xiaojinensis*, during an *Ophiocordyceps sinensis* fungal infection. Insect biochemistry and molecular biology 64, 1-15.

Meng, Y., Omuro, N., Funaguma, S., Daimon, T., Kawaoka, S., Katsuma, S., Shimada, T., 2008. Prominent down-regulation of storage protein genes after bacterial challenge in eri-silkworm, *Samia cynthia ricini*. Archives of insect biochemistry and physiology 67, 9-19.

Michel, T., Reichhart, J.M., Hoffmann, J.A., Royet, J., 2001. *Drosophila* Toll is activated by Gram-positive bacteria through a circulating peptidoglycan recognition protein. Nature 414, 756-759.

Middleton, C.A., Nongthomba, U., Parry, K., Sweeney, S.T., Sparrow, J.C., Elliott, C.J., 2006. Neuromuscular organization and aminergic modulation of contractions in the *Drosophila* ovary. BMC biology 4, 17.

Monner, D.A., Jonsson, S., Boman, H.G., 1971. Ampicillin-resistant mutants of *Escherichia coli* K-12 with lipopolysaccharide alterations affecting mating ability and susceptibility to sex-specific bacteriophages. Journal of bacteriology 107, 420-432.

Montenegro, H., Solferini, V.N., Klaczko, L.B., Hurst, G.D., 2005. Male-killing *Spiroplasma* naturally infecting *Drosophila melanogaster*. Insect molecular biology 14, 281-287.

Montllor, C.B., Maxmen, A., Purcell, A.H., 2002. Facultative bacterial endosymbionts benefit pea aphids *Acyrthosiphon pisum* under heat stress. Ecol. Entomol. 27, 189–195.

Moran, N.A., McCutcheon, J.P., Nakabachi, A., 2008. Genomics and evolution of heritable bacterial symbionts. Annual review of genetics 42, 165-190.

References

Moya, A., Pereto, J., Gil, R., Latorre, A., 2008. Learning how to live together: genomic insights into prokaryote-animal symbioses. Nature reviews. Genetics 9, 218-229.

Murphy, C.I., Piwnica-Worms, H., Grunwald, S., Romanow, W.G., Francis, N., Fan, H.Y., 2004. Expression and purification of recombinant proteins using the baculovirus system. Current protocols in molecular biology / edited by Frederick M. Ausubel ... [et al.] Chapter 16, Unit 16 11.

Myllymaki, H., Valanne, S., Ramet, M., 2014. The *Drosophila* IMD Signaling Pathway. Journal of immunology 192, 3455-3462.

Mylonakis, E., Podsiadlowski, L., Muhammed, M., Vilcinskas, A., 2016. Diversity, evolution and medical applications of insect antimicrobial peptides. Philosophical transactions of the Royal Society of London. Series B, Biological sciences 371.

Nair, D.G., Fry, B.G., Alewood, P., Kumar, P.P., Kini, R.M., 2007. Antimicrobial activity of omwaprin, a new member of the waprin family of snake venom proteins. The Biochemical journal 402, 93-104.

Nakabachi, A., Shigenobu, S., Sakazume, N., Shiraki, T., Hayashizaki, Y., Carninci, P., Ishikawa, H., Kudo, T., Fukatsu, T., 2005. Transcriptome analysis of the aphid bacteriocyte, the symbiotic host cell that harbors an endocellular mutualistic bacterium, *Buchnera*. Proceedings of the National Academy of Sciences of the United States of America 102, 5477-5482.

Nappi, A.J., Vass, E., 1993. Melanogenesis and the generation of cytotoxic molecules during insect cellular immune reactions. Pigment cell research / sponsored by the European Society for Pigment Cell Research and the International Pigment Cell Society 6, 117-126.

Neuenkirchen, N., 2012. An *in vitro* system for the biogenesis of small nuclear ribonucleoprotein particles, Chair of Biochemistry. University of Würzburg.

Nielsen, H., Engelbrecht, J., Brunak, S., von Heijne, G., 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein engineering 10, 1-6.

Niere, M., Dettloff, M., Maier, T., Ziegler, M., Wiesner, A., 2001. Insect immune activation by apolipophorin III is correlated with the lipid-binding properties of this protein. Biochemistry 40, 11502-11508.

Nishikori, K., Kubo, T., Morioka, M., 2009a. Morph-dependent expression and subcellular localization of host serine carboxypeptidase in bacteriocytes of the pea aphid associated with degradation of the endosymbiotic bacterium *Buchnera*. Zoological science 26, 415-420.

Nishikori, K., Morioka, K., Kubo, T., Morioka, M., 2009b. Age- and morph-dependent activation of the lysosomal system and *Buchnera* degradation in aphid endosymbiosis. Journal of insect physiology 55, 351-357.

Nogge, G., 1981. Significance of symbionts for the maintenance of an optimal nutritional state for successful reproduction in hematophagous arthropods.

Noselli, S., Agnes, F., 1999. Roles of the JNK signaling pathway in *Drosophila* morphogenesis. Current opinion in genetics & development 9, 466-472.

References

Novy, R., Drott, D., Yaeger, K., Mierendorf, R., 2001. Overcoming the codon bias of *E. coli* for enhanced protein expression. inNovations 12.

Nygaard, S., Zhang, G., Schiott, M., Li, C., Wurm, Y., Hu, H., Zhou, J., Ji, L., Qiu, F., Rasmussen, M., Pan, H., Hauser, F., Krogh, A., Grimmelikhuijzen, C.J., Wang, J., Boomsma, J.J., 2011. The genome of the leaf-cutting ant *Acromyrmex echinatior* suggests key adaptations to advanced social life and fungus farming. Genome research 21, 1339-1348.

O'Reilly, D.R., Miller, L.K., Luckow, V.A., 1993. Baculovirus Expression Vectors: A Laboratory Manual. Oxford, Oxford University Press.

Oakeson, K.F., Gil, R., Clayton, A.L., Dunn, D.M., von Niederhausern, A.C., Hamil, C., Aoyagi, A., Duval, B., Baca, A., Silva, F.J., Vallier, A., Jackson, D.G., Latorre, A., Weiss, R.B., Heddi, A., Moya, A., Dale, C., 2014. Genome degeneration and adaptation in a nascent stage of symbiosis. Genome biology and evolution 6, 76-93.

Ochiai, M., Ashida, M., 2000. A pattern-recognition protein for beta-1,3-glucan. The binding domain and the cDNA cloning of beta-1,3-glucan recognition protein from the silkworm, *Bombyx mori*. The Journal of biological chemistry 275, 4995-5002.

Ogino, H., Wachi, M., Ishii, A., Iwai, N., Nishida, T., Yamada, S., Nagai, K., Sugai, M., 2004. FtsZ-dependent localization of GroEL protein at possible division sites. Genes to cells : devoted to molecular & cellular mechanisms 9, 765-771.

Oliver, K.M., Degnan, P.H., Burke, G.R., Moran, N.A., 2010. Facultative symbionts in aphids and the horizontal transfer of ecologically important traits. Annu Rev Entomol 55, 247-266.

Oliver, K.M., Moran, N.A., Hunter, M.S., 2005. Variation in resistance to parasitism in aphids is due to symbionts not host genotype. Proceedings of the National Academy of Sciences of the United States of America 102, 12795-12800.

Oliver, K.M., Russell, J.A., Moran, N.A., Hunter, M.S., 2003. Facultative bacterial symbionts in aphids confer resistance to parasitic wasps. Proceedings of the National Academy of Sciences of the United States of America 100, 1803-1807.

Ortius-Lechner, D., Maile, R., Morgan, E.D., Boomsma, J.J., 2000. Metapleural gland secretion of the leaf-cutter ant *Acromyrmex octospinosus*: New compounds and their functional significance. J. Chem. Ecol. 26, 1667–1683.

Otti, O., Tragust, S., Feldhaar, H., 2014. Unifying external and internal immune defences. Trends in ecology & evolution 29, 625-634.

Oxley, P.R., Ji, L., Fetter-Pruneda, I., McKenzie, S.K., Li, C., Hu, H., Zhang, G., Kronauer, D.J., 2014. The genome of the clonal raider ant *Cerapachys biroi*. Current biology : CB 24, 451-458.

Paredes, J.C., Welchman, D.P., Poidevin, M., Lemaitre, B., 2011. Negative regulation by amidase PGRPs shapes the *Drosophila* antibacterial response and protects the fly from innocuous infection. Immunity 35, 770-779.

Park, J.M., Brady, H., Ruocco, M.G., Sun, H., Williams, D., Lee, S.J., Kato, T., Jr., Richards, N., Chan, K., Mercurio, F., Karin, M., Wasserman, S.A., 2004. Targeting of TAK1 by the NF-kappa B

References

protein Relish regulates the JNK-mediated immune response in *Drosophila*. Genes & development 18, 584-594.

Pauls, D., Chen, J., Reiher, W., Vanselow, J.T., Schlosser, A., Kahnt, J., Wegener, C., 2014. Peptidomics and processing of regulatory peptides in the fruit fly *Drosophila*. EuPA Open Proteomics 3, 114−127.

Pfaffl, M.W., Tichopad, A., Prgomet, C., Neuvians, T.P., 2004. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper--Excel-based tool using pair-wise correlations. Biotechnology letters 26, 509-515.

Pham, D.Q., Winzerling, J.J., 2010. Insect ferritins: Typical or atypical? Biochimica et biophysica acta 1800, 824-833.

Phan, J., Zdanov, A., Evdokimov, A.G., Tropea, J.E., Peters, H.K., 3rd, Kapust, R.B., Li, M., Wlodawer, A., Waugh, D.S., 2002. Structural basis for the substrate specificity of tobacco etch virus protease. The Journal of biological chemistry 277, 50564-50572.

Philips, J.A., Rubin, E.J., Perrimon, N., 2005. *Drosophila* RNAi screen reveals CD36 family member required for mycobacterial infection. Science 309, 1251-1253.

Poulsen, J.K., Bot, A.N.M., Nielsen, M.G., Boomsma, J.J., 2002. Experimental evidence for the costs and hygienic significance of the antibiotic metapleural gland secretion in leaf-cutting ants. Behav Ecol Sociobiol 52, 151–157.

Pratt, C.C., Weers, P.M., 2004. Lipopolysaccharide binding of an exchangeable apolipoprotein, apolipophorin III, from *Galleria mellonella*. Biological chemistry 385, 1113-1119.

Purcell, J., Chapuisat, M., 2014. Foster carers influence brood pathogen resistance in ants. Proceedings. Biological sciences / The Royal Society 281.

Qiu, P., Pan, P.C., Govind, S., 1998. A role for the *Drosophila* Toll/Cactus pathway in larval hematopoiesis. Development 125, 1909-1920.

Rahnamaeian, M., Cytrynska, M., Zdybicka-Barabas, A., Dobslaff, K., Wiesner, J., Twyman, R.M., Zuchner, T., Sadd, B.M., Regoes, R.R., Schmid-Hempel, P., Vilcinskas, A., 2015. Insect antimicrobial peptides show potentiating functional interactions against Gram-negative bacteria. Proceedings. Biological sciences / The Royal Society 282, 20150293.

Rämet, M., Lanot, R., Zachary, D., Manfruelli, P., 2002a. JNK signaling pathway is required for efficient wound healing in *Drosophila*. Developmental biology 241, 145-156.

Rämet, M., Manfruelli, P., Pearson, A., Mathey-Prevot, B., Ezekowitz, R.A., 2002b. Functional genomic analysis of phagocytosis and identification of a *Drosophila* receptor for *E. coli*. Nature 416, 644-648.

Ramsey, J.S., Rider, D.S., Walsh, T.K., De Vos, M., Gordon, K.H., Ponnala, L., Macmil, S.L., Roe, B.A., Jander, G., 2010. Comparative analysis of detoxification enzymes in *Acyrthosiphon pisum* and *Myzus persicae*. Insect molecular biology 19 Suppl 2, 155-164.

References

Randolt, K., Gimple, O., Geissendorfer, J., Reinders, J., Prusko, C., Mueller, M.J., Albert, S., Tautz, J., Beier, H., 2008. Immune-related proteins induced in the hemolymph after aseptic and septic injury differ in honey bee worker larvae and adults. Archives of insect biochemistry and physiology 69, 155-167.

Ranson, H., Claudianos, C., Ortelli, F., Abgrall, C., Hemingway, J., Sharakhova, M.V., Unger, M.F., Collins, F.H., Feyereisen, R., 2002. Evolution of supergene families associated with insecticide resistance. Science 298, 179-181.

Rappsilber, J., Ishihama, Y., Mann, M., 2003. Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. Analytical chemistry 75, 663-670.

Ratzka, C., 2012. Immune responses of the ant *Camponotus floridanus* towards pathogens and its obligate mutualistic endosymbiont *Blochmannia floridanus*. PhD Thesis, University Würzburg.

Ratzka, C., Forster, F., Liang, C., Kupper, M., Dandekar, T., Feldhaar, H., Gross, R., 2012a. Molecular characterization of antimicrobial peptide genes of the carpenter ant *Camponotus floridanus*. PloS one 7, e43036.

Ratzka, C., Gross, R., Feldhaar, H., 2012b. Endosymbiont Tolerance and Control within Insect Hosts. Insects 3, 553-572.

Ratzka, C., Gross, R., Feldhaar, H., 2013a. Gene expression analysis of the endosymbiont-bearing midgut tissue during ontogeny of the carpenter ant *Camponotus floridanus*. Journal of insect physiology 59, 611-623.

Ratzka, C., Gross, R., Feldhaar, H., 2013b. Systemic gene knockdown in *Camponotus floridanus* workers by feeding of dsRNA. Insectes Sociaux 60, 475-484.

Ratzka, C., Liang, C., Dandekar, T., Gross, R., Feldhaar, H., 2011. Immune response of the ant *Camponotus floridanus* against pathogens and its obligate mutualistic endosymbiont. Insect biochemistry and molecular biology 41, 529-536.

Reddy, K.V., Yedery, R.D., Aranha, C., 2004. Antimicrobial peptides: premises and promises. International journal of antimicrobial agents 24, 536-547.

Reed, L.J., Muench, H., 1938. A simple method of estimating fifty percent endpoints. The American Journal of Hygiene 27, 493-497.

Reynolds, E.S., 1963. The use of lead citrate at high pH as an electron-opaque stain in electron microscopy. The Journal of cell biology 17, 208-212.

Riddell, C., Adams, S., Schmid-Hempel, P., Mallon, E.B., 2009. Differential expression of immune defences is associated with specific host-parasite interactions in insects. PloS one 4, e7621.

Romanelli, A., Moggio, L., Montella, R.C., Campiglia, P., Iannaccone, M., Capuano, F., Pedone, C., Capparelli, R., 2011. Peptides from Royal Jelly: studies on the antimicrobial activity of jelleins, jelleins analogs and synergy with temporins. Journal of peptide science : an official publication of the European Peptide Society 17, 348-352.

References

Rosengaus, R.B., Guldin, M.R., Traniello, J.F.A., 1998. Inhibitory effect of termite fecal pellets on fungal spore germination. J. Chem. Ecol. 24, 1697–1706.

Rosengaus, R.B., Malak, T., Mackintosh, C., 2013. Immune-priming in ant larvae: social immunity does not undermine individual immunity. Biology letters 9, 20130563.

Royet, J., Gupta, D., Dziarski, R., 2011. Peptidoglycan recognition proteins: modulators of the microbiome and inflammation. Nature reviews. Immunology 11, 837-851.

Rozen, S., Skaletsky, H., 2000. Primer3 on the WWW for general users and for biologist programmers. Methods in molecular biology 132, 365-386.

Russell, J.A., Latorre, A., Sabater-Munoz, B., Moya, A., Moran, N.A., 2003. Side-stepping secondary symbionts: widespread horizontal transfer across and beyond the *Aphidoidea*. Molecular ecology 12, 1061-1075.

Ryu, J.H., Kim, S.H., Lee, H.Y., Bai, J.Y., Nam, Y.D., Bae, J.W., Lee, D.G., Shin, S.C., Ha, E.M., Lee, W.J., 2008. Innate immune homeostasis by the homeobox gene caudal and commensal-gut mutualism in *Drosophila*. Science 319, 777-782.

Sacchi, L., Grigolo, A., Mazzini, M., Bigliardi, E., Baccetti, B., Laudani, U., 1988. Symbionts in the oocytes of *Blattella germanica* (L.) (Dictyoptera: Blattellidae): Their mode of transmission. Int. J. Insect Morphol. & Embryol. 17, 437-446.

Sackton, T.B., Werren, J.H., Clark, A.G., 2013. Characterizing the infection-induced transcriptome of *Nasonia vitripennis* reveals a preponderance of taxonomically-restricted immune genes. PloS one 8, e83984.

Sakaguchi, B., Poulson, D.F., 1961. Distribution of "sex-ratio" agent in tissues of *Drosophila willistoni*. Genetics 46, 1665-1676.

Salmela, H., Amdam, G.V., Freitak, D., 2015. Transfer of Immunity from Mother to Offspring Is Mediated via Egg-Yolk Protein Vitellogenin. PLoS pathogens 11, e1005015.

Sameshima, S., Hasegawa, E., Kitade, O., Minaka, N., Matsumoto, T., 1999. Phylogenetic comparison of endosymbionts with their host ants based on molecular evidence. Zoological Sci 16, 993-1000.

Sauer, C., Dudaczek, D., Holldobler, B., Gross, R., 2002. Tissue localization of the endosymbiotic bacterium "*Candidatus* Blochmannia floridanus" in adults and larvae of the carpenter ant *Camponotus floridanus*. Applied and environmental microbiology 68, 4187-4193.

Sauer, C., Stackebrandt, E., Gadau, J., Holldobler, B., Gross, R., 2000. Systematic relationships and cospeciation of bacterial endosymbionts and their carpenter ant host species: proposal of the new taxon *Candidatus* Blochmannia gen. nov. International journal of systematic and evolutionary microbiology 50 Pt 5, 1877-1886.

Schägger, H., 2006. Tricine-SDS-PAGE. Nature protocols 1, 16-22.

Schägger, H., von Jagow, G., 1987. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. Analytical biochemistry 166, 368-379.

References

Schmitt, F., Vanselow, J.T., Schlosser, A., Kahnt, J., Rossler, W., Wegener, C., 2015. Neuropeptidomics of the carpenter ant *Camponotus floridanus*. Journal of proteome research 14, 1504-1514.

Schröder, D., Deppisch, H., Obermayer, M., Krohne, G., Stackebrandt, E., Hölldobler, B., Goebel, W., Gross, R., 1996. Intracellular endosymbiotic bacteria of *Camponotus* species (carpenter ants): systematics, evolution and ultrastructural characterization. Molecular microbiology 21, 479-489.

Serebrov, V.V., Alekseev, A.A., Glupov, V.V., 2001. Changes in the activity and pattern of hemolymph esterases in the larvae of wax moth *Galleria mellonella* L. (Lepidoptera, Pyralidae) during mycosis. Izvestiia Akademii nauk. Seriia biologicheskaia / Rossiiskaia akademiia nauk, 588-592.

Shai, Y., 1999. Mechanism of the binding, insertion and destabilization of phospholipid bilayer membranes by alpha-helical antimicrobial and cell non-selective membrane-lytic peptides. Biochimica et biophysica acta 1462, 55-70.

Sharp, P.M., Li, W.-H., 1986. The codon adaptation index - a measure of directional synonymous codon usage bias, and its potential applications. Nucleic acids research 15, 1281-1295.

Shi, L., Paskewitz, S.M., 2006. Proteomics and insect immunity. ISJ 3, 4-17.

Shi, X.Z., Zhong, X., Yu, X.Q., 2012. *Drosophila melanogaster* NPC2 proteins bind bacterial cell wall components and may function in immune signal pathways. Insect biochemistry and molecular biology 42, 545-556.

Shiotsuki, T., Kato, Y., 1999. Induction of carboxylesterase isozymes in *Bombyx mori* by *E. coli* infection. Insect biochemistry and molecular biology 29, 731-736.

Silverman, N., Zhou, R., Erlich, R.L., Hunter, M., Bernstein, E., Schneider, D., Maniatis, T., 2003. Immune activation of NF-kappaB and JNK requires *Drosophila* TAK1. The Journal of biological chemistry 278, 48928-48934.

Simiczyjew, B., 2003. Germ Cell Cluster Differentiation in Polytrophic Ovarioles of Hanging-Flies (Mecoptera: Bittacidae). Zoologica Poloniae 48, 71-79.

Sluss, H.K., Han, Z., Barrett, T., Goberdhan, D.C., Wilson, C., Davis, R.J., Ip, Y.T., 1996. A JNK signal transduction pathway that mediates morphogenesis and an immune response in *Drosophila*. Genes & development 10, 2745-2758.

Smith, C.D., Zimin, A., Holt, C., Abouheif, E., Benton, R., Cash, E., Croset, V., Currie, C.R., Elhaik, E., Elsik, C.G., Fave, M.J., Fernandes, V., Gadau, J., Gibson, J.D., Graur, D., Grubbs, K.J., Hagen, D.E., Helmkampf, M., Holley, J.A., Hu, H., Viniegra, A.S., Johnson, B.R., Johnson, R.M., Khila, A., Kim, J.W., Laird, J., Mathis, K.A., Moeller, J.A., Munoz-Torres, M.C., Murphy, M.C., Nakamura, R., Nigam, S., Overson, R.P., Placek, J.E., Rajakumar, R., Reese, J.T., Robertson, H.M., Smith, C.R., Suarez, A.V., Suen, G., Suhr, E.L., Tao, S., Torres, C.W., van Wilgenburg, E., Viljakainen, L., Walden, K.K., Wild, A.L., Yandell, M., Yorke, J.A., Tsutsui, N.D., 2011a. Draft genome of the globally widespread and invasive Argentine ant (*Linepithema humile*). Proceedings of the National Academy of Sciences of the United States of America 108, 5673-5678.

Smith, C.R., Smith, C.D., Robertson, H.M., Helmkampf, M., Zimin, A., Yandell, M., Holt, C., Hu, H., Abouheif, E., Benton, R., Cash, E., Croset, V., Currie, C.R., Elhaik, E., Elsik, C.G., Fave, M.J.,

References

Fernandes, V., Gibson, J.D., Graur, D., Gronenberg, W., Grubbs, K.J., Hagen, D.E., Viniegra, A.S., Johnson, B.R., Johnson, R.M., Khila, A., Kim, J.W., Mathis, K.A., Munoz-Torres, M.C., Murphy, M.C., Mustard, J.A., Nakamura, R., Niehuis, O., Nigam, S., Overson, R.P., Placek, J.E., Rajakumar, R., Reese, J.T., Suen, G., Tao, S., Torres, C.W., Tsutsui, N.D., Viljakainen, L., Wolschin, F., Gadau, J., 2011b. Draft genome of the red harvester ant *Pogonomyrmex barbatus*. Proceedings of the National Academy of Sciences of the United States of America 108, 5667-5672.

Smith, G.E., Ju, G., Ericson, B.L., Moschera, J., Lahm, H.W., Chizzonite, R., Summers, M.D., 1985. Modification and secretion of human interleukin 2 produced in insect cells by a baculovirus expression vector. Proceedings of the National Academy of Sciences of the United States of America 82, 8404-8408.

Smith, V.J., Fernandes, J.M., Kemp, G.D., Hauton, C., 2008. Crustins: enigmatic WAP domain-containing antibacterial proteins from crustaceans. Developmental and comparative immunology 32, 758-772.

Söderhäll, K., Cerenius, L., 1998. Role of the prophenoloxidase-activating system in invertebrate immunity. Current opinion in immunology 10, 23-28.

Soloviev, M., Finch, P., 2006. Peptidomics: bridging the gap between proteome and metabolome. Proteomics 6, 744-747.

Southey, B.R., Rodriguez-Zas, S.L., Sweedler, J.V., 2006. Prediction of neuropeptide prohormone cleavages with application to RFamides. Peptides 27, 1087-1098.

Stanke, M., Keller, O., Gunduz, I., Hayes, A., Waack, S., Morgenstern, B., 2006a. AUGUSTUS: *ab initio* prediction of alternative transcripts. Nucleic acids research 34, W435-439.

Stanke, M., Schoffmann, O., Morgenstern, B., Waack, S., 2006b. Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. BMC bioinformatics 7, 62.

Stanke, M., Waack, S., 2003. Gene prediction with a hidden Markov model and a new intron submodel. Bioinformatics 19 Suppl 2, ii215-225.

Steiner, H., 2004. Peptidoglycan recognition proteins: on and off switches for innate immunity. Immunological reviews 198, 83-96.

Steiner, H., Hultmark, D., Engstrom, A., Bennich, H., Boman, H.G., 1981. Sequence and specificity of two antibacterial proteins involved in insect immunity. Nature 292, 246-248.

Stoll, S., 2009. Funktionelle Analyse von *Blochmannia floridanus*, dem primären Endosymbionten der Rossameise *Camponotus floridanus*. PhD Thesis, Universität Würzburg.

Stoll, S., Feldhaar, H., Fraunholz, M.J., Gross, R., 2010. Bacteriocyte dynamics during development of a holometabolous insect, the carpenter ant *Camponotus floridanus*. BMC microbiology 10, 308.

Stoll, S., Feldhaar, H., Gross, R., 2009. Transcriptional profiling of the endosymbiont *Blochmannia floridanus* during different developmental stages of its holometabolous ant host. Environmental microbiology 11, 877-888.

References

Stow, A., Beattie, A., 2008. Chemical and genetic defenses against disease in insect societies. Brain, behavior, and immunity 22, 1009-1013.

Strand, M.R., 2008. The insect cellular immune response. Insect Science 15, 1-14.

Stroschein-Stevenson, S.L., Foley, E., O'Farrell, P.H., Johnson, A.D., 2006. Identification of *Drosophila* gene products required for phagocytosis of *Candida albicans*. Plos Biol 4, e4.

Suen, G., Teiling, C., Li, L., Holt, C., Abouheif, E., Bornberg-Bauer, E., Bouffard, P., Caldera, E.J., Cash, E., Cavanaugh, A., Denas, O., Elhaik, E., Fave, M.J., Gadau, J., Gibson, J.D., Graur, D., Grubbs, K.J., Hagen, D.E., Harkins, T.T., Helmkampf, M., Hu, H., Johnson, B.R., Kim, J., Marsh, S.E., Moeller, J.A., Munoz-Torres, M.C., Murphy, M.C., Naughton, M.C., Nigam, S., Overson, R., Rajakumar, R., Reese, J.T., Scott, J.J., Smith, C.R., Tao, S., Tsutsui, N.D., Viljakainen, L., Wissler, L., Yandell, M.D., Zimmer, F., Taylor, J., Slater, S.C., Clifton, S.W., Warren, W.C., Elsik, C.G., Smith, C.D., Weinstock, G.M., Gerardo, N.M., Currie, C.R., 2011. The genome sequence of the leaf-cutter ant *Atta cephalotes* reveals insights into its obligate symbiotic lifestyle. PLoS genetics 7, e1002007.

Suthianthong, P., Pulsook, N., Supungul, P., Tassanakajon, A., Rimphanitchayakit, V., 2011. A double WAP domain-containing protein PmDWD from the black tiger shrimp Penaeus monodon is involved in the controlling of proteinase activities in lymphoid organ. Fish & shellfish immunology 30, 783-790.

Swiatoniowska, M., Ogorzalek, A., Golas, A., Michalik, A., Szklarzewicz, T., 2013. Ultrastructure, distribution, and transovarial transmission of symbiotic microorganisms in *Nysius ericae* and *Nithecus jacobaeae* (Heteroptera: Lygaeidae: Orsillinae). Protoplasma 250, 325-332.

Szklarzewicz, T., Kalandyk-Kolodziejczyk, M., Kot, M., Michalik, A., 2013. Ovary structure and transovarial transmission of endosymbiotic microorganisms in *Marchalina hellenica* (Insecta, Hemiptera, Coccomorpha: Marchalinidae). Acta Zoologica 94, 184-192.

Szklarzewicz, T., Kedra, K., Niznik, S., 2006. Ultrastructure and transovarial transmission of endosymbiotic microorganisms in *Palaeococcus fuscipennis* (Burmeister) (Insecta, Hemiptera, Coccinea: Monophlebidae). Folia Biol (Krakow) 54, 69-74.

Takeda, K., Noguchi, T., Naguro, I., Ichijo, H., 2008. Apoptosis signal-regulating kinase 1 in stress and immune response. Annual review of pharmacology and toxicology 48, 199-225.

Takehana, A., Katsuyama, T., Yano, T., Oshima, Y., Takada, H., Aigaki, T., Kurata, S., 2002. Overexpression of a pattern-recognition receptor, peptidoglycan-recognition protein-LE, activates imd/relish-mediated antibacterial defense and the prophenoloxidase cascade in *Drosophila* larvae. Proceedings of the National Academy of Sciences of the United States of America 99, 13705-13710.

Takehana, A., Yano, T., Mita, S., Kotani, A., Oshima, Y., Kurata, S., 2004. Peptidoglycan recognition protein (PGRP)-LE and PGRP-LC act synergistically in *Drosophila* immunity. The EMBO journal 23, 4690-4700.

Tanaka, E.D., Hartfelder, K., 2004. The initial stages of oogenesis and their relation to differential fertility in the honey bee (*Apis mellifera*) castes. Arthropod structure & development 33, 431-442.

Tang, X., Zhou, B., 2013. Iron homeostasis in insects: Insights from *Drosophila* studies. IUBMB life 65, 863-872.

References

Tian, C., Gao, B., Fang, Q., Ye, G., Zhu, S., 2010. Antimicrobial peptide-like genes in *Nasonia vitripennis*: a genomic perspective. Bmc Genomics 11, 187.

Tong, Z., Li, L., Pawar, R., Zhang, S., 2010. Vitellogenin is an acute phase protein with bacterial-binding and inhibiting activities. Immunobiology 215, 898-902.

Tonk, M., Cabezas-Cruz, A., Valdes, J.J., Rego, R.O., Chrudimska, T., Strnad, M., Sima, R., Bell-Sakyi, L., Franta, Z., Vilcinskas, A., Grubhoffer, L., Rahnamaeian, M., 2014. Defensins from the tick *Ixodes scapularis* are effective against phytopathogenic fungi and the human bacterial pathogen *Listeria grayi*. Parasites & vectors 7, 554.

Tonk, M., Cabezas-Cruz, A., Valdes, J.J., Rego, R.O., Grubhoffer, L., Estrada-Pena, A., Vilcinskas, A., Kotsyfakis, M., Rahnamaeian, M., 2015a. *Ixodes ricinus* defensins attack distantly-related pathogens. Developmental and comparative immunology 53, 358-365.

Tonk, M., Knorr, E., Cabezas-Cruz, A., Valdes, J.J., Kollewe, C., Vilcinskas, A., 2015b. *Tribolium castaneum* defensins are primarily active against Gram-positive bacteria. Journal of invertebrate pathology 132, 208-215.

Tragust, S., Ugelvig, L.V., Chapuisat, M., Heinze, J., Cremer, S., 2013. Pupal cocoons affect sanitary brood care and limit fungal infections in ant colonies. BMC evolutionary biology 13, 225.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., Pachter, L., 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nature protocols 7, 562-578.

Tsuchida, T., Koga, R., Fukatsu, T., 2004. Host plant specialization governed by facultative symbiont. Science 303, 1989.

Valanne, S., Wang, J.H., Ramet, M., 2011. The *Drosophila* Toll signaling pathway. Journal of immunology 186, 649-656.

Vallet-Gely, I., Lemaitre, B., Boccard, F., 2008. Bacterial strategies to overcome insect defences. Nature reviews. Microbiology 6, 302-313.

Vaughn, J.L., Goodwin, R.H., Tompkins, G.J., McCawley, P., 1977. The establishment of two cell lines from the insect *Spodoptera frugiperda* (Lepidoptera; Noctuidae). In vitro 13, 213-217.

Veenstra, J.A., 2000. Mono- and dibasic proteolytic cleavage sites in insect neuroendocrine peptide precursors. Archives of insect biochemistry and physiology 43, 49-63.

Vierstraete, E., Verleyen, P., Baggerman, G., D'Hertog, W., Van den Bergh, G., Arckens, L., De Loof, A., Schoofs, L., 2004a. A proteomic approach for the analysis of instantly released wound and immune proteins in *Drosophila melanogaster* hemolymph. Proceedings of the National Academy of Sciences of the United States of America 101, 470-475.

Vierstraete, E., Verleyen, P., Sas, F., Van den Bergh, G., De Loof, A., Arckens, L., Schoofs, L., 2004b. The instantly released *Drosophila* immune proteome is infection-specific. Biochemical and biophysical research communications 317, 1052-1060.

References

Vigneron, A., Masson, F., Vallier, A., Balmand, S., Rey, M., Vincent-Monegat, C., Aksoy, E., Aubailly-Giraud, E., Zaidman-Remy, A., Heddi, A., 2014. Insects recycle endosymbionts when the benefit is over. Current biology : CB 24, 2267-2273.

Viljakainen, L., Pamilo, P., 2008. Selection on an antimicrobial peptide defensin in ants. Journal of molecular evolution 67, 643-652.

Vlisidou, I., Wood, W., 2015. *Drosophila* blood cells and their role in immune responses. The FEBS journal 282, 1368-1382.

Vogel, H., Altincicek, B., Glockner, G., Vilcinskas, A., 2011. A comprehensive transcriptome and immune-gene repertoire of the lepidopteran model host *Galleria mellonella*. Bmc Genomics 12, 308.

Volz, J., Muller, H.M., Zdanowicz, A., Kafatos, F.C., Osta, M.A., 2006. A genetic module regulates the melanization response of *Anopheles* to *Plasmodium*. Cellular microbiology 8, 1392-1405.

Wang, J., Aksoy, S., 2012. PGRP-LB is a maternally transmitted immune milk protein that influences symbiosis and parasitism in tsetse's offspring. Proceedings of the National Academy of Sciences of the United States of America 109, 10552-10557.

Wang, L., Weber, A.N., Atilano, M.L., Filipe, S.R., Gay, N.J., Ligoxygakis, P., 2006. Sensing of Gram-positive bacteria in *Drosophila*: GNBP1 is needed to process and present peptidoglycan to PGRP-SA. The EMBO journal 25, 5005-5014.

Watson, F.L., Puttmann-Holgado, R., Thomas, F., Lamar, D.L., Hughes, M., Kondo, M., Rebel, V.I., Schmucker, D., 2005. Extensive diversity of Ig-superfamily proteins in the immune system of insects. Science 309, 1874-1878.

Waugh, D.S., 2011. An overview of enzymatic reagents for the removal of affinity tags. Protein expression and purification 80, 283-293.

Weers, P.M., Ryan, R.O., 2006. Apolipophorin III: role model apolipoprotein. Insect biochemistry and molecular biology 36, 231-240.

Weimer, R.M., 2006. Preservation of *C. elegans* tissue via high-pressure freezing and freeze-substitution for ultrastructural analysis and immunocytochemistry. Methods in molecular biology 351, 203-221.

Weis, W.I., Taylor, M.E., Drickamer, K., 1998. The C-type lectin superfamily in the immune system. Immunological reviews 163, 19-34.

Wernegreen, J.J., 2012. Strategies of genomic integration within insect-bacterial mutualisms. The Biological bulletin 223, 112-122.

Wheeler, D.E., Martinez, T., 1995. Storage proteins in ants (Hymenoptera: Formicidae). Comparative biochemistry and physiology. Part B, Biochemistry & molecular biology 112, 15-19.

Whitcomb, R.F., Williamson, D.L., 1975. Helical wall-free prokaryotes in insects: multiplication and pathogenicity. Annals of the New York Academy of Sciences 266, 260-275.

References

Whitten, M.M., Tew, I.F., Lee, B.L., Ratcliffe, N.A., 2004. A novel role for an insect apolipoprotein (apolipophorin III) in beta-1,3-glucan pattern recognition and cellular encapsulation reactions. Journal of immunology 172, 2177-2185.

Wilson-Rich, N., Dres, S.T., Starks, P.T., 2008. The ontogeny of immunity: development of innate immune strength in the honey bee (*Apis mellifera*). Journal of insect physiology 54, 1392-1399.

Wilson, E.O., 1971. The Insect Societies. Belknap Press of Harvard Univ Press, Cambridge, Mass, 548.

Wolschin, F., Holldobler, B., Gross, R., Zientz, E., 2004. Replication of the endosymbiotic bacterium *Blochmannia floridanus* is correlated with the developmental and reproductive stages of its ant host. Applied and environmental microbiology 70, 4096-4102.

Woltedji, D., Fang, Y., Han, B., Feng, M., Li, R., Lu, X., Li, J., 2013. Proteome analysis of hemolymph changes during the larval to pupal development stages of honeybee workers (*Apis mellifera ligustica*). Journal of proteome research 12, 5189-5198.

Wurm, Y., Wang, J., Riba-Grognuz, O., Corona, M., Nygaard, S., Hunt, B.G., Ingram, K.K., Falquet, L., Nipitwattanaphon, M., Gotzek, D., Dijkstra, M.B., Oettler, J., Comtesse, F., Shih, C.J., Wu, W.J., Yang, C.C., Thomas, J., Beaudoing, E., Pradervand, S., Flegel, V., Cook, E.D., Fabbretti, R., Stockinger, H., Long, L., Farmerie, W.G., Oakey, J., Boomsma, J.J., Pamilo, P., Yi, S.V., Heinze, J., Goodisman, M.A., Farinelli, L., Harshman, K., Hulo, N., Cerutti, L., Xenarios, I., Shoemaker, D., Keller, L., 2011. The genome of the fire ant *Solenopsis invicta*. Proceedings of the National Academy of Sciences of the United States of America 108, 5679-5684.

Xu, X., Jin, F., Yu, X., Ji, S., Wang, J., Cheng, H., Wang, C., Zhang, W., 2007. Expression and purification of a recombinant antibacterial peptide, cecropin, from *Escherichia coli*. Protein expression and purification 53, 293-301.

Yamamoto, S., Sakai, N., Nakamura, H., Fukagawa, H., Fukuda, K., Takagi, T., 2011. INOH: ontology-based highly structured database of signal transduction pathways. Database : the journal of biological databases and curation 2011, bar052.

Yan, R., Small, S., Desplan, C., Dearolf, C.R., Darnell, J.E., Jr., 1996. Identification of a *Stat* gene that functions in *Drosophila* development. Cell 84, 421-430.

Yano, T., Mita, S., Ohmori, H., Oshima, Y., Fujimoto, Y., Ueda, R., Takada, H., Goldman, W.E., Fukase, K., Silverman, N., Yoshimori, T., Kurata, S., 2008. Autophagic control of *Listeria* through intracellular innate immune recognition in *Drosophila*. Nature immunology 9, 908-916.

Yek, S.H., Mueller, U.G., 2011. The metapleural gland of ants. Biological reviews of the Cambridge Philosophical Society 86, 774-791.

Yi, H.Y., Chowdhury, M., Huang, Y.D., Yu, X.Q., 2014. Insect antimicrobial peptides and their applications. Applied microbiology and biotechnology 98, 5807-5822.

Yu, X.Q., Kanost, M.R., 2000. Immulectin-2, a lipopolysaccharide-specific lectin from an insect, *Manduca sexta,* is induced in response to Gram-negative bacteria. The Journal of biological chemistry 275, 37373-37381.

References

Yu, X.Q., Zhu, Y.F., Ma, C., Fabrick, J.A., Kanost, M.R., 2002. Pattern recognition proteins in *Manduca sexta* plasma. Insect biochemistry and molecular biology 32, 1287-1293.

Zaidman-Remy, A., Herve, M., Poidevin, M., Pili-Floury, S., Kim, M.S., Blanot, D., Oh, B.H., Ueda, R., Mengin-Lecreulx, D., Lemaitre, B., 2006. The *Drosophila* amidase PGRP-LB modulates the immune response to bacterial infection. Immunity 24, 463-473.

Zaidman-Remy, A., Poidevin, M., Herve, M., Welchman, D.P., Paredes, J.C., Fahlander, C., Steiner, H., Mengin-Lecreulx, D., Lemaitre, B., 2011. *Drosophila* immunity: analysis of PGRP-SB1 expression, enzymatic activity and function. PloS one 6, e17231.

Zchori-Fein, E., Roush, R.T., Rosen, D., 1998. Distribution of parthenogenesis-inducing symbionts in ovaries and eggs of *Aphytis* (Hymentoptera: Aphelinidae). Current microbiology 36, 1-8.

Zdybicka-Barabas, A., Cytrynska, M., 2011. Involvement of apolipophorin III in antibacterial defense of *Galleria mellonella* larvae. Comparative biochemistry and physiology. Part B, Biochemistry & molecular biology 158, 90-98.

Zdybicka-Barabas, A., Cytryńska, M., 2013. Apolipophorins and insects immune response. Invertebrate Survival Journal 10, 58-68.

Zhang, C.R., Zhang, S., Xia, J., Li, F.F., Xia, W.Q., Liu, S.S., Wang, X.W., 2014a. The immune strategy and stress response of the Mediterranean species of the *Bemisia tabaci* complex to an orally delivered bacterial pathogen. PloS one 9, e94477.

Zhang, G., Ghosh, S., 2002. Negative regulation of toll-like receptor-mediated signaling by Tollip. The Journal of biological chemistry 277, 7059-7065.

Zhang, G., Hussain, M., O'Neill, S.L., Asgari, S., 2013. *Wolbachia* uses a host microRNA to regulate transcripts of a methyltransferase, contributing to dengue virus inhibition in *Aedes aegypti*. Proceedings of the National Academy of Sciences of the United States of America 110, 10276-10281.

Zhang, S., Wang, S., Li, H., Li, L., 2011. Vitellogenin, a multivalent sensor and an antimicrobial effector. The international journal of biochemistry & cell biology 43, 303-305.

Zhang, Y., Dong, Z., Wang, D., Wu, Y., Song, Q., Gu, P., Zhao, P., Xia, Q., 2014b. Proteomics of larval hemolymph in *Bombyx mori* reveals various nutrient-storage and immunity-related proteins. Amino acids 46, 1021-1031.

Zhang, Z., Zhu, S., 2012. Comparative genomics analysis of five families of antimicrobial peptide-like genes in seven ant species. Developmental and comparative immunology 38, 262-274.

Zhou, X., Kaya, H.K., Heungens, K., Goodrich-Blair, H., 2002. Response of ants to a deterrent factor(s) produced by the symbiotic bacteria of entomopathogenic nematodes. Applied and environmental microbiology 68, 6202-6209.

Zhu, Y., Johnson, T.J., Myers, A.A., Kanost, M.R., 2003. Identification by subtractive suppression hybridization of bacteria-induced genes expressed in *Manduca sexta* fat body. Insect biochemistry and molecular biology 33, 541-559.

References

Zientz, E., Beyaert, I., Gross, R., Feldhaar, H., 2006. Relevance of the endosymbiosis of *Blochmannia floridanus* and carpenter ants at different stages of the life cycle of the host. Applied and environmental microbiology 72, 6027-6033.

Zientz, E., Dandekar, T., Gross, R., 2004. Metabolic interdependence of obligate intracellular bacteria and their insect hosts. Microbiology and molecular biology reviews : MMBR 68, 745-770.

Zou, Z., Lopez, D.L., Kanost, M.R., Evans, J.D., Jiang, H., 2006. Comparative analysis of serine protease-related genes in the honey bee genome: possible involvement in embryonic development and innate immunity. Insect molecular biology 15, 603-614.

Appendix

I   Evaluation of baculovirus titres using end-point dilution

Table 19 Determination of baculovirus titre concentrations using end-point dilution. Previous to protein expression in insect cells, the number of infectious virus particles in the virus stock has to be determined. All recombinant baculoviruses produced in this work carry the gene encoding the enhanced green fluorescent protein (EGFP), and successful transfection can be monitored via fluorescence microscopy. For end-point dilution adapted from the protocol of O'Reilly (1993), ten-fold serial dilutions (10^{-1} to 10^{-9}) of the baculovirus titre were prepared and uninfected *Sf*21 cells were diluted to a concentration of 0.5×10^6 cells/ml. Then, 10 µl of the virus dilution and 60 µl of the uninfected cells were mixed and seeded into a 96-well plate. In total, 12 replicates of each virus dilution were added. The 96-well plates were wrapped in plastic foil to prevent evaporation of the culture medium during the incubation at 27°C for 9-10 d. Afterwards, each well was checked for infection and wells containing at least one cell expressing EGFP were scored as positive. The results of the infection were inserted into the excel spreadsheet below and the number of infectious particles per volume was determined by the method of Reed and Muench (1938).

Date:	DD/MM/YYYY		Recombinant Baculovirus:		HisTEVhld,EGFP

Cell type:	Sf21		Stock generation:	P2

Dilutions	Number of Infected Wells	Number of Uninfected Wells	Total Number Infected	Total number Uninfected	% Total infected	Above 0.5	% Above 0.5	% Below 0.5	Log Dilution Above 50%
1.00E-02	12	0	58	0	100.00	true	0.00	0.00	0.00
1.00E-03	12	0	46	0	100.00	true	0.00	0.00	0.00
1.00E-04	12	0	34	0	100.00	true	0.00	0.00	0.00
1.00E-05	12	0	22	0	100.00	true	0.00	0.00	0.00
1.00E-06	10	2	10	2	83.33	true	83.33	0.00	-6.00
1.00E-07	0	12	0	14	0.00	false	0.00	0.00	0.00
1.00E-08	0	12	0	26	0.00	false	0.00	0.00	0.00
1.00E-09	0	12	0	38	0.00	false	0.00	0.00	0.00

Num. wells	12					SUM	83.33	0.00	-6
mls/well	0.01								

Prop. Dist.	0.400
Log TCID	-6.400
TCID50	3.98E-07
1/TCID50	2.51E+06
TCID50/ml	2.51E+08

pfu/ml	1.74E+08	viral titre

II *C. floridanus* **Protein Orthologs with immune-related function shared with *D. melanogaster*, *A. mellifera* and *N. vitripennis***

Table 20 List of *C. floridanus* protein orthologs shared with *D. melanogaster*, *A. mellifera* and *N. vitripennis*. The 65 orthologs comprise proteins involved in microbial recognition and immune signalling pathways, as well as serine proteases. The respective function or immune pathways, as well as the annotations are listed according to their accession number within the new genome annotation of *C. floridanus*.

Accession Number	Function/Immune pathways	Annotation
Cflo_N_g11181t1	Coagulation, Hematopoiesis, Encapsulatin/Nodulation	Hemocytin
Cflo_N_g1329t1	crq family scavanger receptor, Microbial recognition	Scavenger receptor class B member 1 (Fragment)
Cflo_N_g14772t1	crq family scavanger receptor, Microbial recognition	scavenger receptor class B member 1-like
Cflo_N_g15204t1	crq family scavanger receptor, Microbial recognition	Scavenger receptor class B member 1
Cflo_N_g9951t1	crq family scavanger receptor, Microbial recognition	Scavenger receptor class B member 1
Cflo_N_g6152t1	Draper, Microbial recognition, Hemocyte receptors, Phagocytosis, Laminin protein	Multiple epidermal growth factor-like domains 10
Cflo_N_g9696t1	Eater, Hemocyte receptor, Microbial recognition, Phagocytosis	von Willebrand factor D and EGF domain-containing protein
Cflo_N_g9695t1	EGF, Hemocyte receptor, Microbial recognition, Phagocytosis	von Willebrand factor D and EGF domain-containing protein
Cflo_N_g2961t1	Encapsulation lectin, Galectin, Microbial recognition	Macrophage mannose receptor 1 (Fragment)
Cflo_N_g5742t1	GNBP protein, Microbial recognition, Beta-1,3-glucan-binding protein, Toll signal transduction pathway	Beta-1,3-glucan-binding protein (GNBP)
Cflo_N_g15215t1	GNBP protein, Toll signal transduction pathway, Microbial recognition	Beta-1,3-glucan-binding protein (GNBP)
Cflo_N_g12197t1	Hemocyte receptors, crq family scavanger receptor, Phagocytosis, Microbial recognition	Protein croquemort
Cflo_N_g6595t1	Hemocyte receptors, crq family scavanger receptor, Phagocytosis, Microbial recognition	Protein croquemort
Cflo_N_g79t1	Hemocyte receptors, Scavanger family, Phagocytosis, Microbial recognition	MAM domain-containing glycosylphosphatidylinositol anchor protein 1
Cflo_N_g79t2	Hemocyte receptors, Scavanger family, Phagocytosis, Microbial recognition	MAM domain-containing glycosylphosphatidylinositol anchor protein 1
Cflo_N_g10553t1	Imd signal transduction pathway	Ubiquitin-conjugating enzyme E2-17 kDa, putative
Cflo_N_g10862t1	Imd signal transduction pathway	Death domain-containing adapter protein BG4
Cflo_N_g2123t1	Imd signal transduction pathway	Apoptosis 2 inhibitor
Cflo_N_g5881t1	Imd signal transduction pathway	Inhibitor of nuclear factor kappa-B kinase subunit beta
Cflo_N_g7081t1	Imd signal transduction pathway	Receptor-interacting serine/threonine-protein kinase 1
Cflo_N_g9451t1	Imd signal transduction pathway	Mitogen-activated protein kinase kinase kinase 7
Cflo_N_g9451t2	Imd signal transduction pathway	Mitogen-activated protein kinase kinase kinase 7
Cflo_N_g11129t1	Jak-Stat signal transduction pathway	Suppressor of cytokine signaling 5
Cflo_N_g13035t1	Jak-Stat signal transduction pathway	Signal transducer and activator of

		transcription
Cflo_N_g4220t1	Jak-Stat signal transduction pathway	tyrosine-protein kinase hopscotch isoform X4
Cflo_N_g6115t1	Jak-Stat signal transduction pathway	Cytokine receptor
Cflo_N_g15516t1	JNK signal transduction pathway	Dual specificity mitogen-activated protein kinase kinase 7
Cflo_N_g3291t1	JNK signal transduction pathway	Transcription factor AP-1
Cflo_N_g647t1	JNK signal transduction pathway	Dual specificity mitogen-activated protein kinase kinase 7
Cflo_N_g647t2	JNK signal transduction pathway	Dual specificity mitogen-activated protein kinase kinase 7
Cflo_N_g6920t1	JNK signal transduction pathway, MAPK pathway	stress-activated protein kinase JNK
Cflo_N_g6920t2	JNK signal transduction pathway, MAPK pathway	Stress-activated protein kinase JNK
Cflo_N_g11714t1	JNK signal transduction pathway, Rho GTPase cytoskeleton	Ras-related protein Rac1
Cflo_N_g4036t1	c-type Lysozyme	Lysozyme c-1
Cflo_N_g5519t1	c-type Lysozyme	Lysozyme c-1
Cflo_N_g1918t1	Melanization proPO	Phenoloxidase subunit A3
Cflo_N_g10272t1	PGRP protein, Imd signal transduction pathway, Microbial recognition	PGRP-LC
Cflo_N_g102t2	PGRP protein, Imd signal transduction pathway, Microbial recognition	PGRP-SC2
Cflo_N_g102t1	PGRP protein, Microbial recognition	PGRP-SC2
Cflo_N_g8526t1	PGRP proteins, Microbial recognition, Toll signal transduction pathway	PGRP2 (=PGRP-SA)
Cflo_N_g14922t1	RNA-I antiviral defense, Toll signal transduction pathway	Insulin-like growth factor-binding protein complex acid labile chain
Cflo_N_g5858t1	RNA-I antiviral defense, Toll signal transduction pathway	Protein toll (Fragment)
Cflo_N_g6513t1	RNA-I antiviral defense, Toll signal transduction pathway	Protein toll
Cflo_N_g8274t1	RNA-I antiviral defense, Toll signal transduction pathway	Protein toll
Cflo_N_g8278t1	RNA-I antiviral defense, Toll signal transduction pathway	Protein toll
Cflo_N_g13088t1	Serpin, Melanization, Toll signal transduction pathway	Serpin B10
Cflo_N_g13089t1	Serpin, Melanization, Toll signal transduction pathway	Serpin B10
Cflo_N_g13089t2	Serpin, Melanization, Toll signal transduction pathway	Serpin B10
Cflo_N_g13089t3	Serpin, Melanization, Toll signal transduction pathway	Serpin B10
Cflo_N_g10213t1	Serine protease, Melanization, Toll signal transduction pathway	Thiamine transporter 2
Cflo_N_g7438t1	Serine protease, Melanization, Toll signal transduction pathway	Coagulation factor IX (Fragment)
Cflo_N_g8442t1	Serine protease, Melanization, Toll signal transduction pathway	Serine protease persephone
Cflo_N_g8446t1	Serine protease, Melanization, Toll signal transduction pathway	Serine protease persephone
Cflo_N_g9525t2	Serine protease, Melanization, Toll signal transduction pathway	Serine protease easter 2
Cflo_N_g9745t1	TEPs, Phagocytosis	CD109 antigen
Cflo_N_g4492t1	TEPs, Phagocytosis, Microbial recognition	CD109 antigen
Cflo_N_g7345t1	TEPs, Phagocytosis, Microbial recognition	Alpha-2-macroglobulin
Cflo_N_g9745t2	TEPs, Phagocytosis, Microbial recognition	CD109 antigen (Fragment)
Cflo_N_g11593t1	Toll signal transduction pathway	Myeloid differentiation primary

		response protein MyD88
Cflo_N_g12735t1	Toll signal transduction pathway	Protein spätzle
Cflo_N_g1330t1	Toll signal transduction pathway	Serine/threonine-protein kinase pelle
Cflo_N_g14413t1	Toll signal transduction pathway	NF-kappa-B inhibitor cactus
Cflo_N_g14414t1	Toll signal transduction pathway	stress-induced-phosphoprotein 1-like
Cflo_N_g3305t1	Toll signal transduction pathway	Embryonic polarity protein dorsal
Cflo_N_g9743t1	Toll signal transduction pathway	Protein toll

III Statistical analysis of gene expression data after immune challenge

Table 21 Statistical analysis of expression data of several genes in *C. floridanus* larvae. Larvae of six different *C. floridanus* colonies were pricked with a mixture of heat-killed bacteria, and mRNA samples were taken 12 hours post immune challenge. X-fold expression changed were determined using the ddCt method, and significance of results was calculated using a two-sided dependent T-test (p < 0.05 is significant).

Columns L90–L132 under "Normalized value (dCT)" appear first for IMMUNISATION (columns 3–8) and then for CONTROL (columns 11–16).

Gene name	Cflo_New	L90	L96	L152	L79	L264	L132	mean dCT (immune)	Std.Dv.	L90	L96	L152	L79	L264	L132	mean dCT (control)	Std.Dv.	p value (two sided, depend. t-test)	Mean x-fold expression (2^{ddCT})	Std.Dv.
housekeeping																				
rpL32	Cflo_N_g6999t1	-0.05	0.83	-0.05	-0.93	-1.04	-2.50	-0.621417	1.144092	-0.02	0.62	-0.07	-0.91	-1.12	-1.89	-0.565333	0.906342	0.649936	1.06	0.24
genes with suspected immune function																				
hppaci	Cflo_N_g1319t1	-1.05	-1.09	-0.95	-0.03	2.95	1.15	0.16300	1.614011	-2.01	-2.52	-2.13	-3.25	-3.59	-3.18	-2.77992	0.648935	0.021987	0.25	0.22
hpchit	Cflo_N_g2215t2	2.89	5.06	5.00	6.96	8.97	9.13	6.335583	2.465511	1.22	2.68	3.03	2.10	1.92	1.49	2.071583	0.688938	0.010909	0.13	0.14
chymo	Cflo_N_g907t1	-2.03	-1.11	0.03	1.20	5.93	3.77	1.29933	3.037101	-2.91	-2.58	-1.99	-3.54	-2.18	-2.87	-2.67767	0.557658	0.022383	0.20	0.22
chito	Cflo_N_g2277t1	-0.08	-0.09	1.18	1.80	1.93	1.52	1.045333	0.912388	-0.17	-1.10	0.05	-0.52	-0.40	-1.28	-0.570250	0.522348	0.012248	0.41	0.30
scav	Cflo_N_g9950t1	4.89	5.92	5.02	9.92	9.15	9.33	7.370667	2.335419	5.26	5.71	5.77	5.32	5.09	7.05	5.698667	0.713608	0.136726	0.69	0.69
Posub	Cflo_N_g1918t1	6.08	6.80	5.90	7.73	7.89	6.98	6.897167	0.816825	5.10	5.63	4.98	5.20	5.69	5.02	5.268833	0.311397	0.002148	0.35	0.16
hex	Cflo_N_g6501t1	-2.01	-1.29	-2.01	1.95	-0.08	0.30	-0.52492	1.550687	-2.74	-3.31	-3.66	-3.48	-2.48	-2.96	-3.10425	0.451255	0.011587	0.25	0.20
cP45018a1	Cflo_N_g2542t1	7.19	7.89	8.66	7.83	7.00	6.83	7.566833	0.688333	7.25	6.40	7.11	7.49	7.15	6.92	7.053000	0.370657	0.175397	0.78	0.35
zcp	Cflo_N_g8803t1	11.93	12.94	11.99	12.77	11.46	12.55	12.27092	0.569673	10.21	11.27	10.93	12.41	9.68	13.10	11.26517	1.295792	0.046680	0.61	0.46
ltp	Cflo_N_g14858t1	5.99	6.85	7.02	9.73	9.67	9.94	8.201000	1.768390	5.97	6.67	6.08	7.70	7.30	6.98	6.782250	0.681777	0.035266	0.49	0.37
sushi	Cflo_N_g10836t1	10.15	10.98	9.72	12.58	11.84	11.10	11.06142	1.055288	9.89	10.88	10.00	12.32	11.85	10.47	10.90225	0.992789	0.258473	0.91	0.19
MPI	Cflo_N_g622t1	5.53	5.88	3.91	5.03	3.94	4.45	4.788500	0.825045	6.20	6.46	6.01	8.29	7.67	6.83	6.910917	0.895570	0.010317	5.92	4.68
yellow	Cflo_N_g14239t1	5.99	6.21	5.78	6.79	6.94	6.65	6.390917	0.468911	7.24	7.72	7.11	7.55	7.91	7.73	7.543583	0.311478	0.000130	2.25	0.41
SOCS2	Cflo_N_g4920t1	5.88	6.80	4.76	1.97	1.99	1.50	3.815000	2.289712	6.77	7.69	6.03	4.40	4.64	4.92	5.741583	1.311819	0.006696	4.76	3.49
LAP61281	Cflo_N_g14504t1	1.96	2.40	1.89	3.01	2.98	2.51	2.456750	0.481375	2.97	2.73	2.09	2.95	3.53	2.49	2.793000	0.490045	0.095943	1.31	0.40
coactus	Cflo_N_g14413t1	7.42	7.17	5.66	4.81	4.79	5.28	5.855917	1.163652	7.24	7.66	6.10	8.52	8.65	8.96	7.855667	1.077497	0.052278	7.35	6.75
vitel	Cflo_N_g8262t1	11.99	12.73	11.88	11.75	10.93	10.36	11.60733	0.838097	12.28	12.63	11.70	11.51	11.62	10.56	11.71558	0.712995	0.488477	1.11	0.29
transf	Cflo_N_g7714t1	-0.13	-0.12	-1.05	0.95	0.60	0.09	0.055833	0.690092	4.18	3.14	2.90	3.83	4.64	2.04	3.455250	0.946798	0.000230	12.10	6.12
hp70940	Cflo_N_g9484t1	7.56	8.41	6.95	7.04	6.47	6.34	7.129583	0.765968	7.96	8.72	8.09	10.00	9.59	9.83	9.033417	0.897982	0.023627	5.42	4.37
PHR	Cflo_N_g6985t1	6.65	6.88	4.37	4.91	4.39	3.76	5.16025	1.295339	10.27	10.95	10.03	10.33	9.77	9.34	10.11392	0.545917	0.000035	35.28	16.45
ester	Cflo_N_g6748t1	0.92	1.88	0.99	2.03	1.42	1.31	1.425333	0.451277	7.03	5.55	7.03	7.07	6.59	6.07	6.555583	0.627386	0.000034	40.53	22.20
hp671l2	Cflo_N_g597t1	9.54	8.85	6.96	3.97	4.66	4.09	6.34417	2.465589	14.23	15.58	14.68	14.28	15.14	15.43	14.85692	0.535956	0.000502	936.46	1015.38
PGRP-LB	Cflo_N_g105t1	4.75	3.93	3.56	4.14	4.10	4.07	4.090917	0.383125	8.11	6.84	7.24	7.77	8.22	7.64	7.638000	0.526434	0.000004	12.07	6.27
PGRP-SA	Cflo_N_g8526t1	3.13	2.67	2.86	3.83	3.84	3.55	3.312083	0.502559	5.76	5.89	5.50	6.36	6.35	6.61	6.080417	0.426368	0.000003	6.94	1.54
Rel	Cflo_N_g6082t1	5.31	5.42	4.96	4.09	4.08	3.69	4.591167	0.729725	6.00	5.88	5.56	5.46	5.88	5.96	5.790000	0.223595	0.010335	2.57	1.37
Hym	Cflo_N_g14777t1	2.50	1.70	0.35	3.91	3.24	3.11	2.468333	1.276410	10.53	9.73	11.17	8.40	8.49	9.30	9.602500	1.105545	0.000636	409.63	690.03
Def	Cflo_N_g8312t1	0.65	0.03	-1.27	-1.52	-0.24	-1.80	-0.692500	0.975916	3.51	3.44	3.19	1.78	1.50	1.39	2.467583	1.014093	0.000316	10.37	6.27
Lyso i-typ	Cflo_N_g1023t1	4.25	4.42	3.51	4.75	4.67	4.20	4.299833	0.445717	4.39	4.82	4.12	4.64	4.91	4.23	4.519250	0.324800	0.092808	1.18	0.22
TyrOH	Cflo_N_g3222t2	6.60	6.45	3.92	4.33	3.89	4.79	4.996167	1.230038	7.86	9.21	7.68	8.63	8.12	8.22	8.119333	0.330033	0.002033	11.44	7.42
ImdKl	Cflo_N_g7081t1	4.94	5.00	4.83	5.84	6.14	7.19	5.657000	0.921693	6.19	6.07	5.80	7.49	7.13	7.11	6.630833	0.696290	0.009050	2.08	0.71
Dorsal	Cflo_N_g3305t1	6.79	7.01	6.67	6.25	5.85	6.60	6.529583	0.415192	7.15	7.03	6.99	7.75	6.98	7.75	7.274500	0.373710	0.027173	1.79	0.72
CathD	Cflo_N_g172t1	13.09	13.25	12.88	16.03	16.61	15.20	14.51042	1.641736	13.76	13.25	13.18	17.74	14.23	16.09	14.70875	1.825664	0.741054	1.52	1.02
MAPKKK7	Cflo_N_g9451t1	10.48	10.89	10.24	11.95	12.22	11.93	11.28617	0.851739	10.70	10.99	10.89	12.96	12.63	12.73	11.81725	1.057420	0.013694	1.48	0.36
NOS1	Cflo_N_g197t1	2.48	3.55	1.96	3.02	3.25	4.53	3.133167	0.887809	2.42	3.93	3.01	5.26	5.01	6.13	4.293167	1.422009	0.022565	2.57	1.40
TEP 2	Cflo_N_g7345t1	4.17	4.03	3.30	3.29	3.00	4.91	3.782833	0.716239	7.46	6.74	7.04	8.79	8.68	8.19	7.814417	0.862162	0.000524	22.62	20.01
TEP 1	Cflo_N_g4492t1	8.30	8.33	7.79	8.16	7.64	8.50	8.120250	0.335982	8.25	8.17	7.97	8.66	8.37	9.29	8.450917	0.468469	0.098947	1.30	0.55
TEP 3	Cflo_N_g9745t1	5.63	5.68	4.83	7.44	7.42	8.88	6.646000	1.515344	5.23	5.37	5.37	7.96	7.07	8.16	6.526667	1.370923	0.600976	0.97	0.37

Appendix

Table 22 Statistical analysis of expression data of several genes in *C. floridanus* workers. Workers of six different *C. floridanus* colonies were pricked with a mixture of heat-killed bacteria, and mRNA samples were taken 12 hours post immune challenge. X-fold expression changed were determined using the ddCt method, and significance of results was calculated using a two-sided dependent T-test (p < 0.05 is significant).

Gene name	Cflo_New	Normalized value (dCT) IMMUNISATION						mean dCT (immune)	Std.Dv.	Normalized value (dCT) CONTROL						mean dCT (control)	Std.Dv.	p value (two sided, depend. t-test)	Mean x-fold expression (2^{ddCT})	Std.Dv.
		W90	W96	W152	W79	W264	W132			W90	W96	W152	W79	W264	W132					
housekeeping																				
rpL32	Cflo_N_g6999rt1	-0.06	-0.97	-0.54	-0.06	-0.03	0.05	-0.267917	0.403956	-0.30	-1.12	-0.23	-0.01	0.08	-0.05	-0.271667	0.441246	0.964640	1.01	0.14
genes with suspected immune function																				
appaci	Cflo_N_g1319rt1	10.08	8.00	10.06	7.26	8.48	8.00	8.644000	1.169242	8.67	8.72	8.66	6.86	8.90	7.89	8.284333	0.782177	0.370991	0.91	0.51
hpchit	Cflo_N_g2215t2	12.29	10.36	11.44	10.96	10.73	12.44	11.36975	0.845726	12.73	11.11	12.21	12.66	12.11	12.95	12.29400	0.664446	0.006420	2.00	0.76
chymo	Cflo_N_g907t1	10.09	10.98	10.49	10.78	9.87	10.82	10.50683	0.440191	9.91	9.86	10.96	11.13	10.78	11.66	10.71633	0.709422	0.529120	1.28	0.54
chito	Cflo_N_g2277t1	4.96	3.09	4.25	5.01	3.33	4.93	4.260917	0.862878	4.63	3.17	3.85	4.97	4.04	4.70	4.226333	0.667952	0.843687	1.01	0.33
scav	Cflo_N_g9950t1	5.96	3.03	4.45	6.00	7.01	4.69	5.190583	1.416333	3.39	1.78	2.93	3.60	3.67	2.93	3.052333	0.698223	0.001089	0.25	0.12
Pxtub	Cflo_N_g1918t1	7.87	5.90	6.38	7.87	8.86	6.83	7.284250	1.105268	7.98	4.91	6.24	10.95	5.65	5.46	5.879333	1.132965	0.053389	0.52	0.40
hex	Cflo_N_g6501t1	11.72	9.03	10.43	10.64	10.92	8.89	10.27267	1.106515	11.70	9.09	9.94	10.95	9.75	9.81	10.20850	0.944005	0.833415	1.05	0.50
cP45018a1	Cflo_N_g5542t1	9.75	9.02	10.41	9.17	10.56	11.07	9.99550	0.819878	10.96	9.47	11.01	10.60	10.22	10.66	10.48758	0.573842	0.175853	1.57	0.80
zcp	Cflo_N_g880t1	10.61	10.91	11.41	9.62	9.96	10.31	10.46917	0.646881	10.84	10.85	11.02	12.02	10.84	10.76	11.05575	0.479579	0.204541	1.90	1.69
lip	Cflo_N_g14858t1	6.98	5.97	6.43	6.88	7.01	6.92	6.698667	0.414941	6.76	5.79	6.42	7.02	7.04	6.94	6.660917	0.485502	0.524045	0.98	0.09
sushi	Cflo_N_g10836t1	14.62	10.94	13.89	14.52	14.91	14.13	13.83308	1.463363	14.56	11.40	13.61	14.24	13.99	13.90	13.61767	1.130502	0.290408	0.90	0.28
MPI	Cflo_N_g622t1	3.33	3.02	3.48	3.98	3.93	4.25	3.663917	0.463138	2.87	3.01	3.39	3.95	4.00	3.96	3.531333	0.513632	0.171695	0.92	0.12
yellow	Cflo_N_g14239t1	6.01	6.03	6.36	5.45	5.94	6.95	6.122917	0.498530	7.05	6.71	6.85	7.09	7.84	7.91	7.241667	0.511143	0.004155	2.31	0.92
SOCS2	Cflo_N_g4920t1	2.63	2.01	2.42	2.61	2.92	3.52	2.684833	0.509582	2.84	2.62	2.94	4.00	3.90	3.94	3.374250	0.635645	0.011092	1.68	0.54
LAP61281	Cflo_N_g1450t1	-1.02	-3.13	-2.58	-1.36	-2.27	-4.09	-2.40733	1.131763	-0.28	-1.91	-2.20	-1.95	-2.81	-4.52	-2.27867	1.381080	0.696116	1.23	0.67
cactus	Cflo_N_g14413t1	3.45	3.12	4.46	2.51	2.98	4.34	3.476750	0.776246	4.79	4.11	4.80	5.08	4.62	5.93	4.887750	0.602765	0.005495	2.97	1.60
vitel	Cflo_N_g8262t1	6.99	5.11	7.02	9.99	10.16	12.81	8.679500	2.806248	11.80	9.69	7.94	8.77	7.22	12.01	9.571000	1.988050	0.524658	9.18	13.13
transf	Cflo_N_g7714t1	-0.24	-2.05	-1.56	0.43	0.49	0.45	-0.413250	1.123271	-0.10	-0.90	-0.12	0.09	0.89	-0.02	-0.028000	0.569530	0.278825	1.48	0.81
hp70940	Cflo_N_g9484t1	1.52	0.95	1.44	1.97	1.98	2.65	1.750333	0.581015	4.76	3.43	4.95	5.80	4.04	4.97	4.658000	0.823859	0.000175	8.30	4.03
PHR	Cflo_N_g6985t1	1.92	2.54	1.67	2.88	2.92	4.89	2.805583	1.141065	2.32	5.10	4.87	5.05	7.06	6.01	5.067250	1.580112	0.009680	6.77	6.01
ester	Cflo_N_g8597t1	2.79	1.62	2.42	3.02	2.98	3.60	2.738667	0.668890	2.85	3.11	2.92	3.28	4.72	3.96	3.472333	0.730754	0.050393	1.85	0.97
hp67112	Cflo_N_g6748t1	7.95	9.02	10.20	8.90	6.99	6.94	8.333567	1.277485	10.47	8.13	11.03	12.00	7.73	12.20	10.25875	1.916755	0.081492	9.44	14.50
PGRP-LB	Cflo_N_g103t1	3.72	3.54	3.90	3.70	4.14	4.62	3.936667	0.393943	5.35	5.74	5.76	5.46	5.30	5.52	5.522333	0.193501	0.000463	3.14	0.99
PGRP-SA	Cflo_N_g8526t1	0.98	0.79	0.56	0.91	1.57	1.55	1.058583	0.412279	1.63	1.84	1.64	1.95	3.04	1.54	1.939583	0.557614	0.008070	1.93	0.60
Rel	Cflo_N_g6082t1	1.57	1.97	2.12	1.03	1.47	2.34	1.748583	0.482947	3.02	4.00	3.97	3.65	4.19	4.29	3.852750	0.463048	0.000127	4.51	1.53
Hym	Cflo_N_g14777t1	-1.72	-2.20	-2.00	-2.43	-1.83	-1.49	-1.94483	0.339296	-3.05	-0.71	-0.61	-0.65	1.10	-1.26	-0.82733	1.401206	0.002657	17.80	14.27
Def	Cflo_N_g8312t1	-4.01	-4.24	-4.57	-2.43	-2.00	-1.95	-3.19900	1.202908	-3.05	-0.71	-0.61	-0.65	1.31	-1.26	-0.82733	1.401206	0.009150	7.34	5.83
Lyso i-typ	Cflo_N_g1023t1	0.62	0.85	0.61	0.95	2.16	2.33	1.350750	0.772589	1.27	2.24	2.17	2.14	3.78	2.94	2.424083	0.849489	0.003253	2.21	0.75
TyrOH	Cflo_N_g5222t2	1.73	1.77	2.47	0.95	1.55	2.34	1.800667	0.551739	3.92	4.92	4.91	4.06	5.27	4.52	4.601583	0.529781	0.000107	7.55	3.38
ImdX1	Cflo_N_g708t1	4.45	5.72	4.55	4.85	4.99	4.80	4.892167	0.451409	4.30	5.42	3.94	4.62	5.31	4.94	4.754417	0.578858	0.350242	0.93	0.21
Dorsal	Cflo_N_g3305t1	4.16	14.24	4.93	4.33	5.09	4.98	6.288333	3.915386	4.30	11.94	4.34	4.97	5.04	5.35	5.989333	2.942584	0.524327	0.97	0.48
CasbD	Cflo_N_g9172t1	10.76	11.69	10.51	13.37	12.40	11.82	11.75767	1.056888	11.37	12.00	10.05	13.33	13.10	11.82	11.94450	1.200371	0.347387	1.18	0.35
MAPKKK7	Cflo_N_g9445t1	9.09	9.20	9.38	7.99	9.19	8.99	8.971333	0.499155	9.14	8.99	8.86	8.17	9.04	9.36	8.925500	0.407732	0.735971	0.99	0.21
NOS1	Cflo_N_g197t1	2.86	13.12	3.76	3.14	3.92	3.66	5.074000	3.959439	3.10	12.09	3.05	3.76	3.76	3.45	4.867750	3.551976	0.439522	0.93	0.38
TEP 2	Cflo_N_g7345t1	1.94	1.83	1.89	1.85	2.90	3.63	2.339333	0.754935	1.71	1.85	1.07	1.71	3.06	2.84	2.039000	0.755721	0.135394	0.84	0.23
TEP 1	Cflo_N_g4492t1	6.99	7.96	7.62	6.24	6.90	7.12	7.138083	0.598843	7.57	6.90	7.08	6.53	7.10	7.75	7.155250	0.445712	0.952707	1.10	0.43
TEP 3	Cflo_N_g9745t1	1.98	13.79	2.79	1.78	2.88	2.61	4.304667	4.668328	2.62	13.13	2.29	2.43	3.08	2.63	4.364417	4.304224	0.802130	1.11	0.40

Table 23 The table lists the results of factorial ANOVAs of relative gene expression levels (dCt-values) in *C. floridanus* for selected genes with "developmental stage" (larva or adult worker) and "treatment" (immune-challenged and untreated) as categorical predictors. df: degree of freedom; SS: sum of squares; MS: mean square. P-values $p < 0.05$ are significant (bold).

gene name		SS	df	MS	F	p
hppaci	Intercept	307.2250	1	307.2250	245.5307	**< 0.0001**
	developmental stage	573.0252	1	573.0252	457.9552	**< 0.0001**
	treatment	16.3606	1	16.3606	13.0752	**0.001723**
	dev. Stage * treatment	10.0098	1	10.0098	7.9997	**0.010384**
	Error	25.0254	20	1.2513		
hpchit	Intercept	1542.816	1	1542.816	800.4107	**< 0.0001**
	developmental stage	349.145	1	349.145	181.1360	**< 0.0001**
	treatment	16.731	1	16.731	8.6800	**0.007985**
	dev. Stage * treatment	40.377	1	40.377	20.9475	**0.000183**
	Error	38.551	20	1.928		
chymo	Intercept	590.7261	1	590.7261	230.9325	**< 0.0001**
	developmental stage	766.2417	1	766.2417	299.5468	**< 0.0001**
	treatment	21.2911	1	21.2911	8.3233	**0.009155**
	dev. Stage * treatment	26.2902	1	26.2902	10.2776	**0.004436**
	Error	51.1601	20	2.5580		
chito	Intercept	120.4851	1	120.4851	209.9027	**< 0.0001**
	developmental stage	96.2922	1	96.2922	167.7551	**< 0.0001**
	treatment	4.0846	1	4.0846	7.1159	**0.014789**
	dev. Stage * treatment	3.7493	1	3.7493	6.5319	**0.018842**
	Error	11.4801	20	0.5740		
scav	Intercept	681.3180	1	681.3180	322.2531	**< 0.0001**
	developmental stage	34.9414	1	34.9414	16.5268	**0.000604**
	treatment	21.7770	1	21.7770	10.3002	**0.004401**
	dev. Stage * treatment	0.3261	1	0.3261	0.1542	0.698677
	Error	42.2847	20	2.1142		
Posub	Intercept	962.3817	1	962.3817	1177.442	**< 0.0001**
	developmental stage	1.4928	1	1.4928	1.826	0.191650
	treatment	13.8009	1	13.8009	16.885	**0.000545**
	dev. Stage * treatment	0.0749	1	0.0749	0.092	0.765273
	Error	16.3470	20	0.8173		
hex	Intercept	425.9849	1	425.9849	360.7152	**< 0.0001**
	developmental stage	871.9623	1	871.9623	738.3597	**< 0.0001**
	treatment	10.4821	1	10.4821	8.8761	**0.007414**
	dev. Stage * treatment	9.4891	1	9.4891	8.0352	**0.010240**
	Error	23.6189	20	1.1809		
P45018a1	Intercept	1848.322	1	1848.322	4584.464	**< 0.0001**
	developmental stage	51.567	1	51.567	127.902	**< 0.0001**
	treatment	0.001	1	0.001	0.002	0.966952
	dev. Stage * treatment	1.518	1	1.518	3.765	0.066577
	Error	8.063	20	0.403		
zcp	Intercept	3045.741	1	3045.741	4593.780	**< 0.0001**
	developmental stage	6.067	1	6.067	9.151	**0.006688**
	treatment	0.264	1	0.264	0.398	0.535519
	dev. Stage * treatment	3.803	1	3.803	5.736	**0.026523**
	Error	13.260	20	0.663		
lip	Intercept	1204.974	1	1204.974	1205.001	**< 0.0001**
	developmental stage	3.954	1	3.954	3.955	0.060603
	treatment	3.182	1	3.182	3.182	0.089628
	dev. Stage * treatment	2.861	1	2.861	2.861	0.106287
	Error	20.000	20	1.000		

gene name		SS	df	MS	F	p
sushi	Intercept	3662.677	1	3662.677	2654.725	**< 0.0001**
	developmental stage	45.162	1	45.162	32.734	**< 0.0001**
	treatment	0.210	1	0.210	0.153	0.700240
	dev. Stage * treatment	0.005	1	0.005	0.003	0.953812
	Error	27.594	20	1.380		
MPI	Intercept	535.5126	1	535.5126	1092.292	**< 0.0001**
	developmental stage	30.4313	1	30.4313	62.071	**< 0.0001**
	treatment	5.9392	1	5.9392	12.114	**0.002359**
	dev. Stage * treatment	7.6275	1	7.6275	15.558	**0.000801**
	Error	9.8053	20	0.4903		
yellow	Intercept	1117.860	1	1117.860	5408.820	**< 0.0001**
	developmental stage	0.487	1	0.487	2.357	0.140361
	treatment	7.739	1	7.739	37.446	**< 0.0001**
	dev. Stage * treatment	0.002	1	0.002	0.008	0.928105
	Error	4.133	20	0.207		
SOCS2	Intercept	365.7736	1	365.7736	191.8216	**< 0.0001**
	developmental stage	18.3488	1	18.3488	9.6226	**0.005619**
	treatment	10.2652	1	10.2652	5.3833	**0.031019**
	dev. Stage * treatment	2.2959	1	2.2959	1.2040	0.285556
	Error	38.1368	20	1.9068		
LAP61281	Intercept	0.4767	1	0.4767	0.5210	0.478775
	developmental stage	148.0787	1	148.0787	161.8287	**< 0.0001**
	treatment	0.3242	1	0.3242	0.3543	0.558350
	dev. Stage * treatment	0.0646	1	0.0646	0.0706	0.793130
	Error	18.3007	20	0.9150		
cact1	Intercept	731.0302	1	731.0302	840.0305	**< 0.0001**
	developmental stage	42.8870	1	42.8870	49.2816	**< 0.0001**
	treatment	17.4498	1	17.4498	20.0517	**0.000230**
	dev. Stage * treatment	0.5199	1	0.5199	0.5975	0.448586
	Error	17.4048	20	0.8702		
vitel	Intercept	2592.523	1	2592.523	795.3661	**< 0.0001**
	developmental stage	38.594	1	38.594	11.8404	**0.002585**
	treatment	1.499	1	1.499	0.4600	0.505415
	dev. Stage * treatment	0.920	1	0.920	0.2823	0.601039
	Error	65.191	20	3.260		
transf	Intercept	14.13582	1	14.13582	19.11049	**0.000295**
	developmental stage	23.43141	1	23.43141	31.67739	**< 0.0001**
	treatment	21.48555	1	21.48555	29.04675	**< 0.0001**
	dev. Stage * treatment	13.62780	1	13.62780	18.42370	**0.000355**
	Error	14.79378	20	0.73969		
hp70940	Intercept	764.1976	1	764.1976	1268.693	**< 0.0001**
	developmental stage	142.7303	1	142.7303	236.956	**< 0.0001**
	treatment	34.7258	1	34.7258	57.651	**< 0.0001**
	dev. Stage * treatment	1.5115	1	1.5115	2.509	0.128856
	Error	12.0470	20	0.6024		
PHR	Intercept	803.6754	1	803.6754	556.6862	**< 0.0001**
	developmental stage	82.1696	1	82.1696	56.9169	**< 0.0001**
	treatment	78.0916	1	78.0916	54.0921	**< 0.0001**
	dev. Stage * treatment	10.8703	1	10.8703	7.5296	**0.012510**
	Error	28.8736	20	1.4437		
ester	Intercept	302.1370	1	302.1370	765.5442	**< 0.0001**
	developmental stage	4.6963	1	4.6963	11.8992	**0.002534**
	treatment	51.5871	1	51.5871	130.7095	**< 0.0001**
	dev. Stage * treatment	28.9883	1	28.9883	73.4496	**< 0.0001**
	Error	7.8934	20	0.3947		

gene name		SS	df	MS	F	p
hp67112	Intercept	2375.523	1	2375.523	814.0722	**< 0.0001**
	developmental stage	10.192	1	10.192	3.4927	0.076356
	treatment	163.360	1	163.360	55.9821	**< 0.0001**
	dev. Stage * treatment	65.136	1	65.136	22.3214	**0.000130**
	Error	58.361	20	2.918		
PGRP-LB	Intercept	673.3917	1	673.3917	4368.730	**< 0.0001**
	developmental stage	7.7288	1	7.7288	50.142	**< 0.0001**
	treatment	39.5177	1	39.5177	256.377	**< 0.0001**
	dev. Stage * treatment	5.7707	1	5.7707	37.439	**< 0.0001**
	Error	3.0828	20	0.1541		
PGRP-2	Intercept	230.2929	1	230.2929	1006.456	**< 0.0001**
	developmental stage	61.3312	1	61.3312	268.038	**< 0.0001**
	treatment	19.9765	1	19.9765	87.304	**< 0.0001**
	dev. Stage * treatment	5.3430	1	5.3430	23.351	**0.000101**
	Error	4.5763	20	0.2288		
Rel	Intercept	383.1605	1	383.1605	1487.793	**< 0.0001**
	developmental stage	34.2702	1	34.2702	133.070	**< 0.0001**
	treatment	16.3647	1	16.3647	63.543	**< 0.0001**
	dev. Stage * treatment	1.2294	1	1.2294	4.774	**0.040967**
	Error	5.1507	20	0.2575		
Hym	Intercept	207.1319	1	207.1319	164.6342	**< 0.0001**
	developmental stage	230.2898	1	230.2898	183.0408	**< 0.0001**
	treatment	171.8661	1	171.8661	136.6040	**< 0.0001**
	dev. Stage * treatment	19.0558	1	19.0558	15.1461	**0.000906**
	Error	25.1627	20	1.2581		
Def	Intercept	7.60219	1	7.60219	5.64048	**0.027666**
	developmental stage	50.48465	1	50.48465	37.45735	**< 0.0001**
	treatment	45.90039	1	45.90039	34.05603	**< 0.0001**
	dev. Stage * treatment	0.93240	1	0.93240	0.69180	0.415374
	Error	26.95581	20	1.34779		
Lyso i-typ	Intercept	237.9101	1	237.9101	586.4609	**< 0.0001**
	developmental stage	38.1667	1	38.1667	94.0829	**< 0.0001**
	treatment	2.5068	1	2.5068	6.1794	**0.021885**
	dev. Stage * treatment	1.0938	1	1.0938	2.6962	0.116221
	Error	8.1134	20	0.4057		
TyrOH	Intercept	571.4138	1	571.4138	1035.640	**< 0.0001**
	developmental stage	67.6016	1	67.6016	122.522	**< 0.0001**
	treatment	52.6421	1	52.6421	95.409	**< 0.0001**
	dev. Stage * treatment	0.1558	1	0.1558	0.282	0.601039
	Error	11.0350	20	0.5517		
ImdK1	Intercept	721.6780	1	721.6780	1541.072	**< 0.0001**
	developmental stage	10.4643	1	10.4643	22.345	**0.000129**
	treatment	1.0486	1	1.0486	2.239	0.150176
	dev. Stage * treatment	1.8534	1	1.8534	3.958	0.060506
	Error	9.3659	20	0.4683		
Dorsal	Intercept	1020.445	1	1020.445	167.9669	**< 0.0001**
	developmental stage	3.494	1	3.494	0.5751	0.457091
	treatment	0.299	1	0.299	0.0491	0.826798
	dev. Stage * treatment	1.632	1	1.632	0.2687	0.609907
	Error	121.505	20	6.075		
CathD	Intercept	4201.001	1	4201.001	1957.325	**< 0.0001**
	developmental stage	45.656	1	45.656	21.272	**0.000169**
	treatment	0.223	1	0.223	0.104	0.750798
	dev. Stage * treatment	0.000	1	0.000	0.000	0.992425
	Error	42.926	20	2.146		

gene name		SS	df	MS	F	p
MAPKKK7	Intercept	2521.531	1	2521.531	4464.866	**< 0.0001**
	developmental stage	40.663	1	40.663	72.001	**< 0.0001**
	treatment	0.353	1	0.353	0.625	0.438321
	dev. Stage * treatment	0.499	1	0.499	0.884	0.358321
	Error	11.295	20	0.565		
NOS1	Intercept	452.4755	1	452.4755	58.18871	**< 0.0001**
	developmental stage	9.4910	1	9.4910	1.22055	0.282367
	treatment	1.3645	1	1.3645	0.17547	0.679760
	dev. Stage * treatment	2.8000	1	2.8000	0.36008	0.555202
	Error	155.5200	20	7.7760		
TEP2	Intercept	382.8289	1	382.8289	638.7501	**< 0.0001**
	developmental stage	78.1691	1	78.1691	130.4252	**< 0.0001**
	treatment	20.8833	1	20.8833	34.8439	**< 0.0001**
	dev. Stage * treatment	28.1483	1	28.1483	46.9654	**< 0.0001**
	Error	11.9868	20	0.5993		
TEP1	Intercept	1428.926	1	1428.926	6424.888	**< 0.0001**
	developmental stage	7.783	1	7.783	34.994	**< 0.0001**
	treatment	0.181	1	0.181	0.816	0.377105
	dev. Stage * treatment	0.147	1	0.147	0.663	0.425140
	Error	4.448	20	0.222		
TEP3	Intercept	715.5931	1	715.5931	64.32973	**< 0.0001**
	developmental stage	30.4234	1	30.4234	2.73497	0.113784
	treatment	0.0053	1	0.0053	0.00048	0.982761
	dev. Stage * treatment	0.0481	1	0.0481	0.00432	0.948220
	Error	222.4766	20	11.1238		

Appendix

Table 24 Tukey's HSD post hoc test of relative gene expression levels (dCt-values) in larvae and workers of *C. floridanus* for putative immune genes with fixed factors: developmental stage (larvae (L) or worker (W)), and treatment (control or injected with heat-killed bacteria). Numbers in bold printing indicate significant differences in gene expression ($p < 0.05$). A p-value > 0.05 indicates, that two samples belong into the same group (compare with Figure 14 and Figure 15).

gene name	dev. stage	treatment	L/c	L/i	W/c	W/i
hppaci	L	control		**0.001143**	**0.000175**	**0.000175**
	L	injected	**0.001143**		**0.000175**	**0.000175**
	W	control	**0.000175**	**0.000175**		0.943639
	W	injected	**0.000175**	**0.000175**	0.943639	
hpchit	L	control		**0.000337**	**0.000175**	**0.000175**
	L	injected	**0.000337**		**0.000176**	**0.000190**
	W	control	**0.000175**	**0.000176**		0.662191
	W	injected	**0.000175**	**0.000190**	0.662191	
chymo	L	control		**0.001922**	**0.000175**	**0.000175**
	L	injected	**0.001922**		**0.000175**	**0.000175**
	W	control	**0.000175**	**0.000175**		0.995810
	W	injected	**0.000175**	**0.000175**	0.995810	
chito	L	control		**0.007339**	**0.000175**	**0.000175**
	L	injected	**0.007339**		**0.000176**	**0.000176**
	W	control	**0.000175**	**0.000176**		0.999834
	W	injected	**0.000175**	**0.000176**	0.999834	
scav	L	control		0.224296	**0.023901**	0.929295
	L	injected	0.224296		**0.000419**	0.075071
	W	control	**0.023901**	**0.000419**		0.082734
	W	injected	0.929295	0.075071	0.082734	
Posub	L	control		**0.025628**	0.652271	**0.005065**
	L	injected	**0.025628**		0.239908	0.879196
	W	control	0.652271	0.239908		0.062244
	W	injected	**0.005065**	0.879196	0.062244	
hex	L	control		**0.002923**	**0.000175**	**0.000175**
	L	injected	**0.002923**		**0.000175**	**0.000175**
	W	control	**0.000175**	**0.000175**		0.999641
	W	injected	**0.000175**	**0.000175**	0.999641	
cP45018a1	L	control		0.512907	**0.000175**	**0.000176**
	L	injected	0.512907		**0.000176**	**0.000181**
	W	control	**0.000175**	**0.000176**		0.548171
	W	injected	**0.000176**	**0.000181**	0.548171	
zcp	L	control		0.175035	0.969795	0.353131
	L	injected	0.175035		0.076864	**0.005394**
	W	control	0.969795	0.076864		0.605131
	W	injected	0.353131	**0.005394**	0.605131	
lip	L	control		0.098271	0.996679	0.998936
	L	injected	0.098271		0.065299	0.074305
	W	control	0.996679	0.065299		0.999906
	W	injected	0.998936	0.074305	0.999906	
sushi	L	control		0.995368	**0.003699**	**0.001863**
	L	injected	0.995368		**0.006205**	**0.003080**
	W	control	**0.003699**	**0.006205**		0.988631
	W	injected	**0.001863**	**0.003080**	0.988631	
MPI	L	control		**0.000365**	**0.000175**	**0.000176**
	L	injected	**0.000365**		**0.026167**	0.051869
	W	control	**0.000175**	**0.026167**		0.987520
	W	injected	**0.000176**	0.051869	0.987520	
yellow	L	control		**0.001605**	0.663846	**0.000306**
	L	injected	**0.001605**		**0.019739**	0.739341
	W	control	0.663846	**0.019739**		**0.002112**
	W	injected	**0.000306**	0.739341	**0.002112**	

gene name	dev. stage	treatment	L/c	L/i	W/c	W/i
SOCS2	L	control		0.106153	**0.035207**	**0.005376**
	L	injected	0.106153		0.944764	0.503553
	W	control	**0.035207**	0.944764		0.822808
	W	injected	**0.005376**	0.503553	0.822808	
LAP61281	L	control		0.928150	**0.000175**	**0.000175**
	L	injected	0.928150		**0.000175**	**0.000175**
	W	control	**0.000175**	**0.000175**		0.995468
	W	injected	**0.000175**	**0.000175**	0.995468	
cact1	L	control		**0.007031**	**0.000280**	**0.000175**
	L	injected	**0.007031**		0.303616	**0.001522**
	W	control	**0.000280**	0.303616		0.071767
	W	injected	**0.000175**	**0.001522**	0.071767	
vitel	L	control		0.999625	0.201228	**0.039623**
	L	injected	0.999625		0.238534	**0.049097**
	W	control	0.201228	0.238534		0.827486
	W	injected	**0.039623**	**0.049097**	0.827486	
transf	L	control		**0.000179**	**0.000177**	**0.000176**
	L	injected	**0.000179**		0.998294	0.781514
	W	control	**0.000177**	0.998294		0.864500
	W	injected	**0.000176**	0.781514	0.864500	
hp70940	L	control		**0.002174**	**0.000175**	**0.000175**
	L	injected	**0.002174**		**0.000279**	**0.000175**
	W	control	**0.000175**	**0.000279**		**0.000184**
	W	injected	**0.000175**	**0.000175**	**0.000184**	
PHR	L	control		**0.000177**	**0.000176**	**0.000175**
	L	injected	**0.000177**		0.999166	**0.014151**
	W	control	**0.000176**	0.999166		**0.018948**
	W	injected	**0.000175**	**0.014151**	**0.018948**	
ester	L	control		**0.000175**	**0.000175**	**0.000175**
	L	injected	**0.000175**		**0.000253**	**0.008609**
	W	control	**0.000175**	**0.000253**		0.212679
	W	injected	**0.000175**	**0.008609**	0.212679	
hp67112	L	control		**0.000175**	**0.000930**	**0.000181**
	L	injected	**0.000175**		**0.003993**	0.214393
	W	control	**0.000930**	**0.003993**		0.239947
	W	injected	**0.000181**	0.214393	0.239947	
PGRP-LB	L	control		**0.000175**	**0.000175**	**0.000175**
	L	injected	**0.000175**		**0.000188**	0.903349
	W	control	**0.000175**	**0.000188**		**0.000178**
	W	injected	**0.000175**	0.903349	**0.000178**	
PGRP-SA	L	control		**0.000175**	**0.000175**	**0.000175**
	L	injected	**0.000175**		**0.000541**	**0.000175**
	W	control	**0.000175**	**0.000541**		**0.022046**
	W	injected	**0.000175**	**0.000175**	**0.022046**	
Rel	L	control		**0.003049**	**0.000181**	**0.000175**
	L	injected	**0.003049**		0.087137	**0.000175**
	W	control	**0.000181**	0.087137		**0.000177**
	W	injected	**0.000175**	**0.000175**	**0.000177**	
Hym	L	control		**0.000175**	**0.000175**	**0.000175**
	L	injected	**0.000175**		0.572303	**0.000179**
	W	control	**0.000175**	0.572303		**0.000279**
	W	injected	**0.000175**	**0.000179**	**0.000279**	
Def	L	control		**0.000843**	**0.000591**	**0.000175**
	L	injected	**0.000843**		0.997083	**0.006631**
	W	control	**0.000591**	0.997083		**0.010322**
	W	injected	**0.000175**	**0.006631**	**0.010322**	

gene name	dev. stage	treatment	L/c	L/i	W/c	W/i
Lyso i-typ	L	control		0.931961	**0.000245**	**0.000175**
	L	injected	0.931961		**0.000444**	**0.000176**
	W	control	**0.000245**	**0.000444**		**0.039122**
	W	injected	**0.000175**	**0.000176**	**0.039122**	
TyrOH	L	control		**0.000176**	**0.000175**	**0.000175**
	L	injected	**0.000176**		0.794560	**0.000176**
	W	control	**0.000175**	0.794560		**0.000183**
	W	injected	**0.000175**	**0.000176**	**0.000183**	
ImdK1	L	control		0.096892	**0.000791**	**0.001576**
	L	injected	0.096892		0.135430	0.245389
	W	control	**0.000791**	0.135430		0.985082
	W	injected	**0.001576**	0.245389	0.985082	
Dorsal	L	control		0.952546	0.803583	0.898634
	L	injected	0.952546		0.980931	0.998267
	W	control	0.803583	0.980931		0.996698
	W	injected	0.898634	0.998267	0.996698	
CathD	L	control		0.995381	**0.018627**	**0.011504**
	L	injected	0.995381		**0.030762**	**0.019186**
	W	control	**0.018627**	**0.030762**		0.996141
	W	injected	**0.011504**	**0.019186**	0.996141	
MAPKKK7	L	control		0.619468	**0.000181**	**0.000182**
	L	injected	0.619468		**0.000298**	**0.000331**
	W	control	**0.000181**	**0.000298**		0.999605
	W	injected	**0.000182**	**0.000331**	0.999605	
NOS1	L	control		0.887843	0.984037	0.961606
	L	injected	0.887843		0.706934	0.630667
	W	control	0.984037	0.706934		0.999283
	W	injected	0.961606	0.630667	0.999283	
TEP2	L	control		**0.000175**	**0.000175**	**0.000175**
	L	injected	**0.000175**		**0.004636**	**0.020251**
	W	control	**0.000175**	**0.004636**		0.906516
	W	injected	**0.000175**	**0.020251**	0.906516	
TEP1	L	control		0.625269	**0.000777**	**0.000694**
	L	injected	0.625269		**0.010191**	**0.008872**
	W	control	**0.000777**	**0.010191**		0.999916
	W	injected	**0.000694**	**0.008872**	0.999916	
TEP3	L	control		0.999920	0.680162	0.661670
	L	injected	0.999920		0.643101	0.624389
	W	control	0.680162	0.643101		0.999990
	W	injected	0.661670	0.624389	0.999990	

IX Identification of hymenoptaecin peptides via mass spectrometry

Table 25 Mass spectrometry analysis of protein bands. A haemolymph sample from *C. floridanus* injected with heat-killed bacteria was taken 48 hours post injection, and proteins were separated via SDS-PAGE (Figure 25). Protein bands comprising proteins with a molecular weight of 10 to 15 kDa were cut out, and proteins were identified via mass spectrometry. The proteins in band 2-4 were identified as hymenoptaecin (Cflo-HD2). Protein scores greater than 49% are significant ($p < 0.05$).

Spot Index	Accession Number	Protein MW [Da]	Protein Score C.I. %
2	Cflo_HD2	10413.1699	100.00
2	Cflo_HD2	10413.1699	99.76
3	Cflo_HD2	10413.1699	100.00
4	Cflo_HD2	10413.1699	100.00
4	Cflo_HD2	10413.1699	100.00

X   Statistical analysis of immune gene expression data in *C. floridanus* ovaries

Table 26 Shapiro-Wilk *W*-test of normality. A p-value > 0.05 indicates that the gene expression (dCt) within a sample is normally distributed.

Gene	Tissue	W	p
PGRP-LB	residual body	0.95546	0.78417
	midgut	0.91001	0.43646
	ovary	0.9953	0.99812
PGRP-SC2	residual body	0.98157	0.9591
	midgut	0.95403	0.77274
	ovary	0.93928	0.65344
PGRP-SA	residual body	0.92127	0.51454
	midgut	0.84578	0.14545
	ovary	0.98879	0.98602
PGRP-LC	residual body	0.87445	0.24451
	midgut	0.7949	0.5285
	ovary	0.82627	0.9992
GNBP	residual body	0.93263	0.60055
	midgut	0.94087	0.66622
	ovary	0.88379	0.28694
GNBP1_3	residual body	0.9836	0.42569
	midgut	0.99473	0.99754
	ovary	0.88159	0.27644
psh	residual body	0.96816	0.87985
	midgut	0.86057	0.19112
	ovary	0.93056	0.58443
relish	residual body	0.75632	0.02295
	midgut	0.90805	0.4237
	ovary	0.84084	0.13246
tollip	residual body	0.95157	0.75299
	midgut	0.89872	0.36641
	ovary	0.89939	0.37033
TEP2	residual body	0.86475	0.20607
	midgut	0.90628	0.41235
	ovary	0.86824	0.21929

Table 27 Factorial ANOVAs of normalized gene expression levels (dCt-values) for putative immune genes within tissues analysed separately. "Tissue" was included as a fixed effect, whereas "tissue nested in colony" ("Tissue(Colony)") and "Colony" were included as random effects. Bold printing indicates significance of the respective effect or interaction of effects. (df: degree of freedom; SS: sum of squares; MS: mean square)

Gene	Effect	Effect F/ R	SS	Degr. of Freedom	MS	p
PGRP-LB	Intercept	Fixed	828.3038	1	828.3038	<0.0001
	Tissue	**Fixed**	**44.7156**	**2**	**22.3578**	**<0.0001**
	Colony*Tissue	**Fixed**	**8.3442**	**3**	**2.7814**	**0.001060**
	Colony	Random		0		
	Tissue(Colony)	Random		0		
	Error		3.1317	12	0.2610	
PGRP-SC2	Intercept	Fixed	251.0119	1	251.0119	<0.0001
	Tissue	**Fixed**	**85.5377**	**2**	**42.7689**	**<0.0001**
	Colony*Tissue	Fixed	0.2609	3	0.0870	0.442855
	Colony	Random		0		
	Tissue(Colony)	Random		0		
	Error		1.0863	12	0.0905	

Gene	Effect	Effect F/ R	SS	Degr. of Freedom	MS	p
PGRP-SA	Intercept	Fixed	961.4843	1	961.4843	<0.0001
	Tissue	**Fixed**	**50.4975**	**2**	**25.2488**	**<0.0001**
	Colony*Tissue	Fixed	0.6505	3	0.2168	0.081358
	Colony	Random		0		
	Tissue(Colony)	Random		0		
	Error		0.9100	12	0.0758	
PGRP-LC	Intercept	Fixed	660.0717	1	660.0717	<0.0001
	Tissue	**Fixed**	**3.8584**	**2**	**1.9292**	**0.003712**
	Colony*Tissue	**Fixed**	**6.9142**	**3**	**2.3047**	**0.000909**
	Colony	Random		0		
	Tissue(Colony)	Random		0		
	Error		2.5032	12	0.2086	
GNBP	Intercept	Fixed	2318.279	1	2318.279	<0.0001
	Tissue	**Fixed**	**115.505**	**2**	**57.753**	**<0.0001**
	Colony*Tissue	**Fixed**	**2.999**	**3**	**1.000**	**0.048412**
	Colony	Random		0		
	Tissue(Colony)	Random		0		
	Error		3.395	12	0.283	
GNBP1_3	Intercept	Fixed	156.4897	1	156.4897	<0.0001
	Tissue	**Fixed**	**14.8656**	**2**	**7.4328**	**<0.0001**
	Colony*Tissue	Fixed	0.3006	3	0.1002	0.146615
	Colony	Random		0		
	Tissue(Colony)	Random		0		
	Error		0.5583	12	0.0465	
psh	Intercept	Fixed	1243.442	1	1243.442	<0.0001
	Tissue	**Fixed**	**207.843**	**2**	**103.922**	**<0.0001**
	Colony*Tissue	Fixed	0.329	3	0.110	0.407492
	Colony	Random		0		
	Tissue(Colony)	Random		0		
	Error		1.257	12	0.105	
relish	Intercept	Fixed	450.3134	1	450.3134	<0.0001
	Tissue	**Fixed**	**37.8038**	**2**	**18.9019**	**<0.0001**
	Colony*Tissue	**Fixed**	**1.1836**	**3**	**0.3945**	**0.011626**
	Colony	Random		0		
	Tissue(Colony)	Random		0		
	Error		0.8314	12	0.0693	
tollip	Intercept	Fixed	892.4468	1	892.4468	<0.0001
	Tissue	**Fixed**	**9.9489**	**2**	**4.9745**	**0.000116**
	Colony*Tissue	Fixed	0.5850	3	0.1950	0.503048
	Colony	Random		0		
	Tissue(Colony)	Random		0		
	Error		2.8224	12	0.2352	
TEP2	Intercept	Fixed	953.7480	1	953.7480	<0.0001
	Tissue	**Fixed**	**251.1126**	**2**	**125.5563**	**<0.0001**
	Colony*Tissue	**Fixed**	**5.7936**	**3**	**1.9312**	**0.004630**
	Colony	Random		0		
	Tissue(Colony)	Random		0		
	Error		3.1418	12	0.2618	

Table 28 Significance of expression (dCT) of putative immune genes. The post hoc test "Fisher's least significant difference" (LSD) was performed to distinguish significant differences of expression (dCT) of several genes within three tissues (residual body, midgut and ovary). The gene expression was analysed within tissues from two different *C. floridanus* colonies (C90 and C152).

Gene	Cell #	Colony	Tissue	{1}	{2}	{3}	{4}	{5}	{6}
							p-value		
	1	C90	residual body		0.000138	0.007712	0.000767	<0.0001	<0.0001
	2	C90	midgut	0.000138		0.040416	<0.0001	0.041481	0.762421
PGRP-LB	3	C90	ovary	0.007712	0.040416		<0.0001	0.000633	0.022959
	4	C152	residual body	0.000767	<0.0001	0.000633		<0.0001	<0.0001
	5	C152	midgut	<0.0001	0.041481	<0.0001	<0.0001		0.071935
	6	C152	ovary	<0.0001	0.762421	0.022959	<0.0001	0.071935	
	1	C90	residual body		0.000779	<0.0001	0.520674	0.000280	<0.0001
	2	C90	midgut	0.000779		<0.0001	0.000252	0.560718	<0.0001
PGRP-SC2	3	C90	ovary	<0.0001	<0.0001		<0.0001	<0.0001	0.174298
	4	C152	residual body	0.520674	0.000252	<0.0001		<0.0001	<0.0001
	5	C152	midgut	0.000280	0.560718	<0.0001	<0.0001		<0.0001
	6	C152	ovary	<0.0001	<0.0001	0.174298	<0.0001	<0.0001	
	1	C90	residual body		<0.0001	<0.0001	0.064727	<0.0001	<0.0001
	2	C90	midgut	<0.0001		<0.0001	<0.0001	0.313050	<0.0001
PGRP-SA	3	C90	ovary	<0.0001	<0.0001		<0.0001	<0.0001	0.092828
	4	C152	residual body	0.064727	<0.0001	<0.0001		<0.0001	<0.0001
	5	C152	midgut	<0.0001	0.313050	<0.0001	<0.0001		<0.0001
	6	C152	ovary	<0.0001	<0.0001	0.092828	<0.0001	<0.0001	
	1	C90	residual body		0.022637	0.059372	0.012003	<0.0001	<0.0001
	2	C90	midgut	0.022637		0.604741	0.738012	0.006082	0.008779
PGRP-LC	3	C90	ovary	0.059372	0.604741		0.399329	0.002293	0.003289
	4	C152	residual body	0.012003	0.738012	0.399329		0.011483	0.016580
	5	C152	midgut	<0.0001	0.006082	0.002293	0.011483		0.846482
	6	C152	ovary	<0.0001	0.008779	0.003289	0.016580	0.846482	
	1	C90	residual body		<0.0001	0.058841	0.933709	<0.0001	0.010527
	2	C90	midgut	<0.0001		<0.0001	<0.0001	0.008915	<0.0001
GNBP	3	C90	ovary	0.058841	<0.0001		0.050578	<0.0001	0.366017
	4	C152	residual body	0.933709	<0.0001	0.050578		<0.0001	0.008990
	5	C152	midgut	<0.0001	0.008915	<0.0001	<0.0001		<0.0001
	6	C152	ovary	0.010527	<0.0001	0.366017	0.008990	<0.0001	
	1	C90	residual body		<0.0001	0.000221	0.054321	<0.0001	<0.0001
	2	C90	midgut	<0.0001		<0.0001	<0.0001	0.277029	<0.0001
GNBP1_3	3	C90	ovary	0.000221	<0.0001		0.009718	<0.0001	0.447061
	4	C152	residual body	0.054321	<0.0001	0.009718		<0.0001	0.002285
	5	C152	midgut	<0.0001	0.277029	<0.0001	<0.0001		<0.0001
	6	C152	ovary	<0.0001	<0.0001	0.447061	0.002285	<0.0001	
	1	C90	residual body		<0.0001	<0.0001	0.102388	<0.0001	<0.0001
	2	C90	midgut	<0.0001		<0.0001	<0.0001	0.947462	<0.0001
psh	3	C90	ovary	<0.0001	<0.0001		<0.0001	<0.0001	0.931412
	4	C152	residual body	0.102388	<0.0001	<0.0001		<0.0001	<0.0001
	5	C152	midgut	<0.0001	0.947462	<0.0001	<0.0001		<0.0001
	6	C152	ovary	<0.0001	<0.0001	0.931412	<0.0001	<0.0001	
	1	C90	residual body		0.046872	<0.0001	0.011149	0.000420	<0.0001
	2	C90	midgut	0.046872		<0.0001	0.449830	0.023072	<0.0001
relish	3	C90	ovary	<0.0001	<0.0001		<0.0001	<0.0001	0.271600
	4	C152	residual body	0.011149	0.449830	<0.0001		0.093390	<0.0001
	5	C152	midgut	0.000420	0.023072	<0.0001	0.093390		<0.0001
	6	C152	ovary	<0.0001	<0.0001	0.271600	<0.0001	<0.0001	
	1	C90	residual body		0.976103	0.002102	0.336894	0.504518	0.000374
	2	C90	midgut	0.976103		0.001989	0.322909	0.486137	0.000355
tollip	3	C90	ovary	0.002102	0.001989		0.013285	0.007437	0.344114
	4	C152	residual body	0.336894	0.322909	0.013285		0.760168	0.002162
	5	C152	midgut	0.504518	0.486137	0.007437	0.760168		0.001234
	6	C152	ovary	0.000374	0.000355	0.344114	0.002162	0.001234	
	1	C90	residual body		<0.0001	0.003399	0.031197	<0.0001	0.707863
	2	C90	midgut	<0.0001		<0.0001	<0.0001	0.967796	<0.0001
TEP2	3	C90	ovary	0.003399	<0.0001		<0.0001	<0.0001	0.001693
	4	C152	residual body	0.031197	<0.0001	<0.0001		<0.0001	0.062260
	5	C152	midgut	<0.0001	0.967796	<0.0001	<0.0001		<0.0001
	6	C152	ovary	0.707863	<0.0001	0.001693	0.062260	<0.0001	

XI   Annual Reviews, Inc - LICENSE, TERMS AND CONDITIONS

This is a License Agreement between Maria Kupper ("You") and Annual Reviews, Inc ("Annual Reviews, Inc") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Annual Reviews, Inc, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	3904311268350
License date	Jul 05, 2016
Licensed content publisher	Annual Reviews, Inc
Licensed content title	Annual Review of Immunology
Licensed content date	Jan 1, 2007
Type of Use	Thesis/Dissertation
Requestor type	Academic institution
Format	Print, Electronic
Portion	chart/graph/table/figure
Number of charts/graphs/tables/figures	1
Title or numeric reference of the portion(s)	page 702, Figure 3
Title of the article or chapter the portion is from	The Host Defense of Drosophila melanogaster
Editor of portion(s)	N/A
Author of portion(s)	Bruno Lemaitre and Jules Hoffmann
Volume of serial or monograph.	25
Page range of the portion	702
Publication date of portion	January 2, 2007
Rights for	Main product
Duration of use	Life of current edition
Creation of copies for the disabled	no
With minor editing privileges	yes
For distribution to	Worldwide
In the following language(s)	Original language of publication
With incidental promotional use	no
The lifetime unit quantity of new product	Up to 499
Made available in the following markets	university
Specified additional information	Minor editing: I would like to remove the lower section showing graphs from the figure. In my thesis I will only show the scheme for the Drosophila immune signalling pathways.
The requesting person/organization is:	Maria Kupper/ Universität Würzburg
Order reference number	
Author/Editor	Maria Kupper

XII   Digital Appendix: List of Content

Table S1: The haemolymph proteome of *C. floridanus*

> Table S1.1: **Original data set.** The table lists proteins identified in the haemolymph of *C. floridanus* (adult workers and larvae). In addition to the original data obtained via mass spectrometry, the respective peptide counts and results of quantification of proteins are presented.

> Table S1.2: List of all proteins (annotation according to Gupta et al., 2015) identified in the haemolymph of *C. floridanus* (adult workers and larvae) and their respective amino acid sequences.

> Table S1.3: List of all proteins which were detected only in adult worker haemolymph

> Table S1.4: List of all proteins which were detected only in larval haemolymph samples

> Table S1.5: List of all proteins which were detected in haemolymph samples of *C. floridanus* workers and larval.

> Table S1.6: List of all proteins which were regulated in the haemolymph of adults upon immune challenge

> Table S1.7: List of all proteins which were regulated in the haemolymph of larvae upon immune challenge

Table S2: Haemolymph proteins and their GO annotations

> Proteins identified in *C. floridanus* haemolymph were assigned to several GO-terms within the categories "biological process" (P), "molecular function" (F) and "cellular compartment" (C). (Analysis performed by Shishir K. Gupta)

Table S3: Haemolymph peptidome of *C. floridanus* larvae

> Table S3.1: **Original list** presenting data from mass spectrometry for the identification of peptides in larval haemolymph.

> Table S3.2: List of all enriched peptides within the larval haemolymph. The table also lists corresponding proteins from which the detected peptides derived. Additionally, the *in silico* predicted peptides are given. Matching amino acid sequences are highlighted in red.

Table S4: Haemolymph peptidome of *C. floridanus* workers

> Table S4.1: **Original list** presenting data from mass spectrometry for the identification of peptides in adult worker haemolymph.

> Table S4.2: List of all enriched peptides within the adult haemolymph. The table also lists corresponding proteins from which the detected peptides derived. Additionally, the *in silico* predicted peptides are given. Matching amino acid sequences are highlighted in red.

Index of Abbreviations

°C	degree Celsius
µg	microgram
µl	microliter
µm	micrometre
µM	micromolar

A

A. mellifera	*Apis mellifera*
A. yanonensis	*Aphytis yanonensis*
A. pisum	*Acyrthosiphon pisum*
aa	amino acids
Acc. No.	Accession number
AMP	Antimicrobial peptide

B

B. aphidicola	*Buchnera aphidicola*
B. floridanus	*Blochmannia floridanus*
BEVS	baculovirus expression vector system
bp	base pairs

C

C-terminus/ C-terminal	carboxyl-terminus; refers to the end of a protein or polypeptide which is terminated by a free carboxyl group (-COOH)
C. floridanus	*Camponotus floridanus*
C. obscurior	*Cardiocondyla obscurior*
C. pennsylvanicus	*Camponotus pennsylvanicus*
CB	cystoblast
cDNA	complementary DNA
CoSO	Cell of Somatic Origin
ColA	Coleoptericin-A

Ct-value	cycle threshold; value in real time PCR indicating the number of PCR cycles required for the fluorescence signal to cross the threshold (exceeding the background fluorescence level)
CV	column volume

D

D. melanogaster	*Drosophila melanogaster*
DAP-type PGN	Diaminopimelic Acid-type Peptidoglycan
DAPI	4´,6-Diamidino-2-phenylindole
dCt-value	DeltaCt-value; the normalised Ct-value is equal to the difference in threshold cycles (Ct) for the target (T) gene and the endogenous reference (R) ($Ct_T - Ct_R$)
ddCt-value	DeltaDeltaCt-value; is equal to the difference in normalised Ct-values (dCt) for treatment (t) and control (c) conditions in an experimental setting ($-(dCt_{T,t} - dCt_{T,c})$)
DNA	Deoxyribonucleic acid
dO	determined oocyte
DOPA	3,4-dihydroxyphenylalanine

E

E. coli	*Escherichia coli*
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid

F

FISH	fluorescence *in situ* hybridisation

G

G. mellonella	*Galleria mellonella*
G. morsitans	*Glossina morsitans*
GNBP	Gram-negative Binding Proteins
GO	Gene Ontology
GSC	germ-line stem cell
GST	Glutathione S-transferase

H

h	hour
His	Histidine
His6/ 6xHis	fusion tag consisting of six histidine residues
hld	hymenoptaecin-like domain
hpi	hours post injection
HPF	High pressure freezing
HPLC	High Performance Liquid Chromatography
hrep	hymenoptaecin repeat domain

I

IMD	immune deficiency

J

JAK/STAT	janus kinase/signal transduction and activator of transcription
JNK	c-jun N-terminal kinase

K

kbp	kilobase pair
kDa	kilo Dalton

L

l	litre
LB	Lysogeny broth
LPS	lipopolysaccharide
LTA	lipoteichoic acid

M

M	Molar
M. luteus	*Micrococcus luteus*
M. sexta	*Manduca sexta*
MAMP	Microbe-Associated Molecular Pattern
mA	milliampere
MCS	multiple cloning site

mg	milligram/ mid gut
min	minute
ml	millilitre
mm	millimetre
mM	millimolar
MOI	multiplicity of infection
mRNA	messenger Ribonucleic acid
MSRO	Melanogaster sex ratio organism

N

N-terminus/ N-terminal	amino-terminus/ NH_2-terminus; refers to the start of a protein or polypeptide terminated by an amino acid with a free amine group
N. vitripennis	*Nasonia vitripennis*
nm	nanometre
nM	nanomolar
NSRO	Nebulosa sex ratio organism

O

OD_{600}	optical density measured at 600 nanometre
ov	ovary

P

p.a.	pro analysi
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PEI	polyethylenimin
PGN	Peptidoglycan
PGRP	Peptidoglycan Recognition Protein
pI	isoelectric point
PO	Phenoloxidase
PPAE	Prophenoloxidase activating enzyme
proPO	Prophenoloxidase
PRR	Pattern recognition receptor

Psh	serine protease Persephone

Q

qRT-PCR	quantitative Reverse transcription PCR

R

rb	residual body parts
RNA	ribonucleic acid
rpm	revolutions per minute
RT	room temperature

S

S. aureus	*Staphylococcus aureus*
S. frugiperda	*Spodoptera frugiperda*
S. glossinidius	*Sodalis glossinidius*
S. nysicola	*Schneideria nysicola*
S. pierantonius	*Sodalis pierantonius*
S. zeamais	*Sitophilus zeamais*
S-tag	Oligopeptide consisting of 15 amino acids derived from the S-peptide of the N-terminus of the subtilisin digested pancreatic ribonuclease A
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
sec/ s	second
Sf21	*Spodoptera frugiperda* 21
SP	Serine Protease
SPE	Spätzle processing enzyme
SZPE	*Sitophilus zeamais* Primary Endosymbiont

T

TEM	Transmission electron microscopy
TEP	thioester-containing protein
TEV protease	*Tobacco Etch Virus nuclear-inclusion-a endopeptidase*
Trx-tag	Thioredoxin-tag
TyrOH	tyrosine hydroxylase

U

  Upd   Unpaired-related Ligand

W

  W. cardiocondylae   *Westeberhardia cardiocondylae*

  W. glossinidia   *Wigglesworthia glossinidia*

X

  x g   times gravity

"Sometimes we do a thing in order to find out the reason for it. Sometimes our actions are questions, not answers."

(John le Carré)